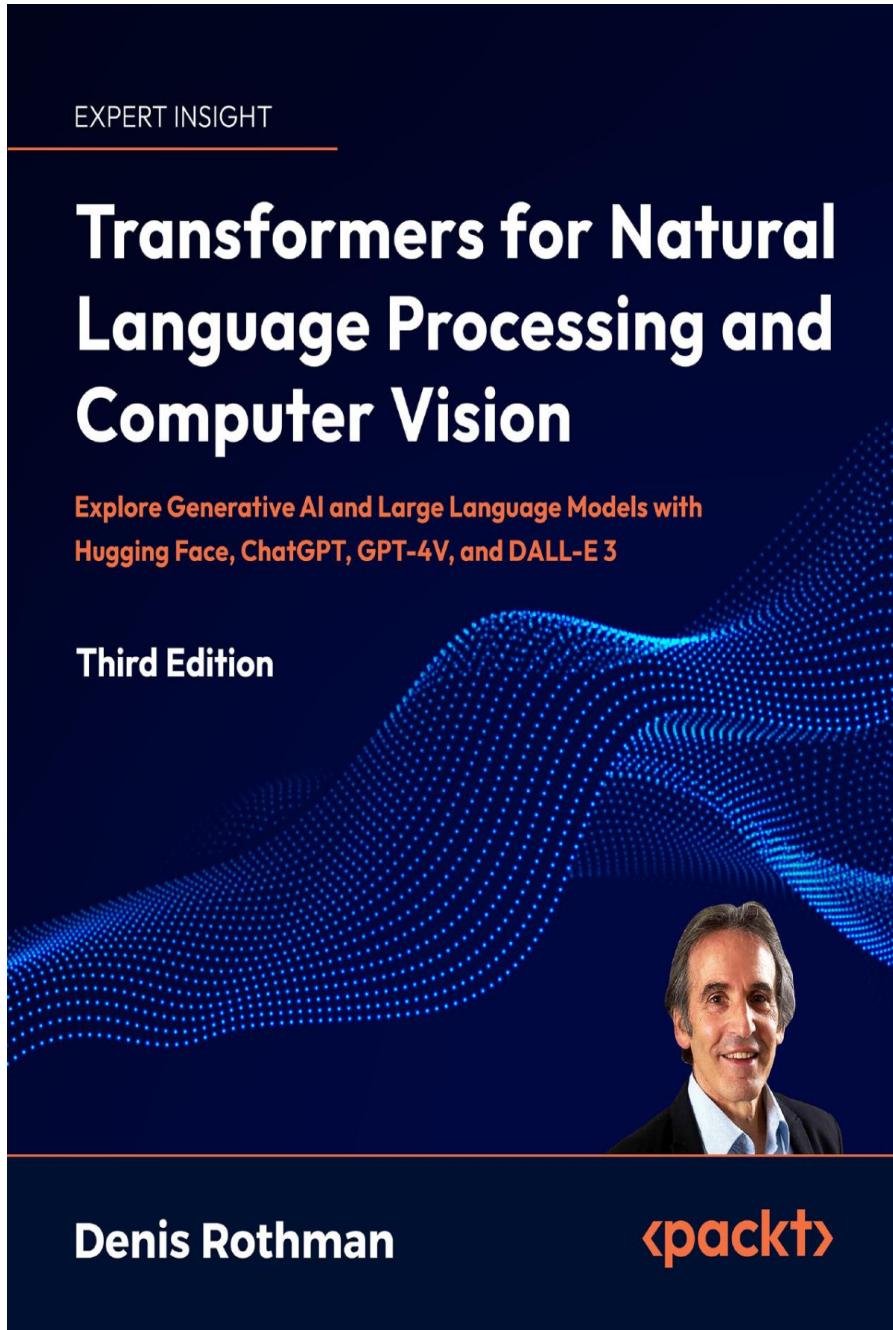


Transformers for Natural Language Processing and Computer Vision, Third Edition Denis Rothman

Visit to download the full and correct content document:
<https://ebookmass.com/product/transformers-for-natural-language-processing-and-computer-vision-third-edition-denis-rothman/>



EXPERT INSIGHT

Transformers for Natural Language Processing and Computer Vision

Explore Generative AI and Large Language Models with
Hugging Face, ChatGPT, GPT-4V, and DALL-E 3

Third Edition



Denis Rothman

<packt>

Transformers for Natural Language Processing and Computer Vision

Third Edition

**Explore Generative AI and Large Language
Models with Hugging Face, ChatGPT, GPT-4V,
and DALL-E 3**

Denis Rothman



BIRMINGHAM—MUMBAI

Transformers for Natural Language Processing and Computer Vision

Third Edition

Copyright © 2024 Packt Publishing

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, without the prior written permission of the publisher, except in the case of brief quotations embedded in critical articles or reviews.

Every effort has been made in the preparation of this book to ensure the accuracy of the information presented. However, the information contained in this book is sold without warranty, either express or implied. Neither the author, nor Packt Publishing or its dealers and distributors, will be held liable for any damages caused or alleged to have been caused directly or indirectly by this book.

Packt Publishing has endeavored to provide trademark information about all of the companies and products mentioned in this book by the appropriate use of capitals. However, Packt Publishing cannot guarantee the accuracy of this information.

Publishing Product Manager: Bhavesh Amin

Acquisition Editor – Peer Reviews: Tejas Mhasvekar

Project Editor: Janice Gonsalves

Content Development Editor: Bhavesh Amin

Copy Editor: Safis Editing

Technical Editor: Karan Sonawane

Proofreader: Safis Editing

Indexer: Rekha Nair

Presentation Designer: Ajay Patule

First published: January 2021

Second edition: February 2022

Third edition: February 2024

Production reference: 1280224

Published by Packt Publishing Ltd.

Grosvenor House

11 St Paul's Square

Birmingham

B3 1RB, UK.

ISBN 978-1-80512-872-4

www.packt.com

Contributors

About the author

Denis Rothman graduated from Sorbonne University and Paris Diderot University, designing one of the first patented encoding and embedding systems. He authored one of the first patented AI cognitive robots and bots. He began his career delivering **Natural Language Processing (NLP)** chatbots for Moët et Chandon and as an **AI tactical defense optimizer** for Airbus (formerly Aerospatiale).

Denis then authored an **AI resource optimizer** for IBM and luxury brands, leading to an **Advanced Planning and Scheduling (APS)** solution used worldwide.

I want to thank the corporations that trusted me from the start to deliver artificial intelligence solutions and shared the risks of continuous innovation. I also want to thank my family, who always believed I would make it.

About the reviewer

George Mihaila has 7 years of research experience with transformer models, having started working with them since they came out in 2017. He is a final-year PhD student in computer science working in research on transformer models in **Natural Language Processing (NLP)**. His research covers both Generative and Predictive NLP modeling.

He has over 6 years of industry experience working in top companies with transformer models and machine learning, covering a broad area from NLP and Computer Vision to Explainability and Causality. George has worked in both science and engineering roles. He is an end-to-end Machine Learning expert leading Research and Development, as well as MLOps, optimization, and deployment.

He was a technical reviewer for the first and second editions of *Transformers for Natural Language Processing* by Denis Rothman.

Join our community on Discord

Join our community's Discord space for discussions with the authors and other readers:

<https://www.packt.link/Transformers>



Contents

Preface

Who this book is for

What this book covers

To get the most out of this book

Get in touch

1. What Are Transformers?

How constant time complexity O(1) changed our lives forever

O(1) attention conquers O(n) recurrent methods

Attention layer

Recurrent layer

The magic of the computational time complexity of an attention layer

Computational time complexity with a CPU

Computational time complexity with a GPU

Computational time complexity with a TPU

TPU-LLM

A brief journey from recurrent to attention

A brief history

From one token to an AI revolution

From one token to everything

Foundation Models

From general purpose to specific tasks

The role of AI professionals

The future of AI professionals

What resources should we use?

Decision-making guidelines

The rise of transformer seamless APIs and assistants

Choosing ready-to-use API-driven libraries

Choosing a cloud platform and transformer model

Summary

Questions

References

Further reading

2. Getting Started with the Architecture of the Transformer Model

The rise of the Transformer: Attention Is All You Need

The encoder stack

Input embedding

Positional encoding

Sublayer 1: Multi-head attention

Sublayer 2: Feedforward network

The decoder stack

Output embedding and position encoding

The attention layers

The FFN sublayer, the post-LN, and the linear layer

Training and performance

Hugging Face transformer models

Summary

Questions

References

Further reading

3. Emergent vs Downstream Tasks: The Unseen Depths of Transformers

The paradigm shift: What is an NLP task?

Inside the head of the attention sublayer of a transformer

Exploring emergence with ChatGPT

Investigating the potential of downstream tasks

Evaluating models with metrics

Accuracy score

F1-score

MCC

Human evaluation

Benchmark tasks and datasets

Defining the SuperGLUE benchmark tasks

Running downstream tasks

The Corpus of Linguistic Acceptability (CoLA)

Stanford Sentiment TreeBank (SST-2)

Microsoft Research Paraphrase Corpus (MRPC)

Winograd schemas

Summary

Questions

References

Further reading

4. Advancements in Translations with Google Trax, Google Translate, and Gemini

Defining machine translation

Human transductions and translations

Machine transductions and translations

Evaluating machine translations

Preprocessing a WMT dataset

Preprocessing the raw data

Finalizing the preprocessing of the datasets

Evaluating machine translations with BLEU

Geometric evaluations

Applying a smoothing technique

Translations with Google Trax

Installing Trax

Creating the Original Transformer model

Initializing the model using pretrained weights

Tokenizing a sentence

Decoding from the Transformer

De-tokenizing and displaying the translation

Translation with Google Translate

Translation with a Google Translate AJAX API Wrapper

Implementing googletrans

Translation with Gemini

Gemini's potential

Summary

Questions

References

Further reading

5. Diving into Fine-Tuning through BERT

The architecture of BERT

The encoder stack

Preparing the pretraining input environment

Pretraining and fine-tuning a BERT model

Fine-tuning BERT

Defining a goal

Hardware constraints

Installing Hugging Face Transformers

[Importing the modules](#)
[Specifying CUDA as the device for torch](#)
[Loading the CoLA dataset](#)
[Creating sentences, label lists, and adding BERT tokens](#)
[Activating the BERT tokenizer](#)
[Processing the data](#)
[Creating attention masks](#)
[Splitting the data into training and validation sets](#)
[Converting all the data into torch tensors](#)
[Selecting a batch size and creating an iterator](#)
[BERT model configuration](#)
[Loading the Hugging Face BERT uncased base model](#)
[Optimizer grouped parameters](#)
[The hyperparameters for the training loop](#)
[The training loop](#)
[Training evaluation](#)
[Predicting and evaluating using the holdout dataset](#)
 [Exploring the prediction process](#)
 [Evaluating using the Matthews correlation coefficient](#)
 [Matthews correlation coefficient evaluation for the whole dataset](#)
[Building a Python interface to interact with the model](#)
 [Saving the model](#)
 [Creating an interface for the trained model](#)
 [Interacting with the model](#)
[Summary](#)
[Questions](#)
[References](#)
[Further reading](#)

6. [Pretraining a Transformer from Scratch through RoBERTa](#)

[Training a tokenizer and pretraining a transformer](#)

[Building KantaiBERT from scratch](#)

[Step 1: Loading the dataset](#)

[Step 2: Installing Hugging Face transformers](#)

[Step 3: Training a tokenizer](#)

[Step 4: Saving the files to disk](#)

[Step 5: Loading the trained tokenizer files](#)

[Step 6: Checking resource constraints: GPU and CUDA](#)
[Step 7: Defining the configuration of the model](#)
[Step 8: Reloading the tokenizer in transformers](#)
[Step 9: Initializing a model from scratch](#)
 [Exploring the parameters](#)
[Step 10: Building the dataset](#)
[Step 11: Defining a data collator](#)
[Step 12: Initializing the trainer](#)
[Step 13: Pretraining the model](#)
[Step 14: Saving the final model \(+tokenizer + config\) to disk](#)
[Step 15: Language modeling with FillMaskPipeline](#)

[Pretraining a Generative AI customer support model on X data](#)

[Step 1: Downloading the dataset](#)
 [Step 2: Installing Hugging Face transformers](#)
 [Step 3: Loading and filtering the data](#)
 [Step 4: Checking Resource Constraints: GPU and CUDA](#)
 [Step 5: Defining the configuration of the model](#)
 [Step 6: Creating and processing the dataset](#)
 [Step 7: Initializing the trainer](#)
 [Step 8: Pretraining the model](#)
 [Step 9: Saving the model](#)
 [Step 10: User interface to chat with the Generative AI agent](#)
 [Further pretraining](#)
 [Limitations](#)

[Next steps](#)

[Summary](#)

[Questions](#)

[References](#)

[Further reading](#)

7. [The Generative AI Revolution with ChatGPT](#)

[GPTs as GPTs](#)

[Improvement](#)

[Diffusion](#)

[New application sectors](#)

[Self-service assistants](#)

[Development assistants](#)

[Pervasiveness](#)

The architecture of OpenAI GPT transformer models

The rise of billion-parameter transformer models

The increasing size of transformer models

Context size and maximum path length

From fine-tuning to zero-shot models

Stacking decoder layers

GPT models

OpenAI models as assistants

ChatGPT provides source code

GitHub Copilot code assistant

General-purpose prompt examples

Getting started with ChatGPT – GPT-4 as an assistant

1. GPT-4 helps to explain how to write source code

2. GPT-4 creates a function to show the YouTube

presentation of GPT-4 by Greg Brockman on March 14, 2023

3. GPT-4 creates an application for WikiArt to display images

4. GPT-4 creates an application to display IMDb reviews

5. GPT-4 creates an application to display a newsfeed

6. GPT-4 creates a k-means clustering (KMC) algorithm

Getting started with the GPT-4 API

Running our first NLP task with GPT-4

Steps 1: Installing OpenAI and Step 2: Entering the API key

Step 3: Running an NLP task with GPT-4

Key hyperparameters

Running multiple NLP tasks

Retrieval Augmented Generation (RAG) with GPT-4

Installation

Document retrieval

Augmented retrieval generation

Summary

Questions

References

Further reading

8. Fine-Tuning OpenAI GPT Models

Risk management

Fine-tuning a GPT model for completion (generative)

1. Preparing the dataset

1.1. Preparing the data in JSON

1.2. Converting the data to JSONL

2. Fine-tuning an original model

3. Running the fine-tuned GPT model

4. Managing fine-tuned jobs and models

Before leaving

Summary

Questions

References

Further reading

9. Shattering the Black Box with Interpretable Tools

Transformer visualization with BertViz

Running BertViz

Step 1: Installing BertViz and importing the modules

Step 2: Load the models and retrieve attention

Step 3: Head view

Step 4: Processing and displaying attention heads

Step 5: Model view

Step 6: Displaying the output probabilities of attention heads

Streaming the output of the attention heads

Visualizing word relationships using attention scores with pandas

exBERT

Interpreting Hugging Face transformers with SHAP

Introducing SHAP

Explaining Hugging Face outputs with SHAP

Transformer visualization via dictionary learning

Transformer factors

Introducing LIME

The visualization interface

Other interpretable AI tools

LIT

[PCA](#)

[Running LIT](#)

[OpenAI LLMs explain neurons in transformers](#)

[Limitations and human control](#)

[Summary](#)

[Questions](#)

[References](#)

[Further reading](#)

[10. Investigating the Role of Tokenizers in Shaping Transformer Models](#)

[Matching datasets and tokenizers](#)

[Best practices](#)

[Step 1: Preprocessing](#)

[Step 2: Quality control](#)

[Step 3: Continuous human quality control](#)

[Word2Vec tokenization](#)

[Case 0: Words in the dataset and the dictionary](#)

[Case 1: Words not in the dataset or the dictionary](#)

[Case 2: Noisy relationships](#)

[Case 3: Words in a text but not in the dictionary](#)

[Case 4: Rare words](#)

[Case 5: Replacing rare words](#)

[Exploring sentence and WordPiece tokenizers to understand the efficiency of subword tokenizers for transformers](#)

[Word and sentence tokenizers](#)

[Sentence tokenization](#)

[Word tokenization](#)

[Regular expression tokenization](#)

[Treebank tokenization](#)

[White space tokenization](#)

[Punkt tokenization](#)

[Word punctuation tokenization](#)

[Multi-word tokenization](#)

[Subword tokenizers](#)

[Unigram language model tokenization](#)

[SentencePiece](#)

[Byte-Pair Encoding \(BPE\)](#)

[WordPiece](#)

Exploring in code

Detecting the type of tokenizer

Displaying token-ID mappings

Analyzing and controlling the quality of token-ID mappings

Summary

Questions

References

Further reading

11. Leveraging LLM Embeddings as an Alternative to Fine-Tuning

LLM embeddings as an alternative to fine-tuning

From prompt design to prompt engineering

Fundamentals of text embedding with NLTK and Gensim

Installing libraries

1. Reading the text file

2. Tokenizing the text with Punkt

Preprocessing the tokens

3. Embedding with Gensim and Word2Vec

4. Model description

5. Accessing a word and vector

6. Exploring Gensim's vector space

7. TensorFlow Projector

Implementing question-answering systems with embedding-based search techniques

1. Installing the libraries and selecting the models

2. Implementing the embedding model and the GPT model

2.1 Evaluating the model with a knowledge base: GPT can answer questions

2.2 Add a knowledge base

2.3 Evaluating the model without a knowledge base: GPT cannot answer questions

3. Prepare search data

4. Search

5. Ask

5.1 Example question

5.2 Troubleshooting wrong answers

Transfer learning with Ada embeddings

1. The Amazon Fine Food Reviews dataset
 - 1.2. Data preparation
2. Running Ada embeddings and saving them for future reuse
3. Clustering
 - 3.1. Find the clusters using k-means clustering
 - 3.2. Display clusters with t-SNE
4. Text samples in the clusters and naming the clusters

Summary

Questions

References

Further reading

12. Toward Syntax-Free Semantic Role Labeling with ChatGPT and GPT-4

Getting started with cutting-edge SRL

Entering the syntax-free world of AI

Defining SRL

Visualizing SRL

SRL experiments with ChatGPT with GPT-4

Basic sample

Difficult sample

Questioning the scope of SRL

The challenges of predicate analysis

Redefining SRL

From task-specific SRL to emergence with ChatGPT

1. Installing OpenAI

2. GPT-4 dialog function

3. SRL

Sample 1 (basic)

Sample 2 (basic)

Sample 3 (basic)

Sample 4 (difficult)

Sample 5 (difficult)

Sample 6 (difficult)

Summary

Questions

References

Further reading

13. Summarization with T5 and ChatGPT

Designing a universal text-to-text model

The rise of text-to-text transformer models

A prefix instead of task-specific formats

The T5 model

Text summarization with T5

Hugging Face

Selecting a Hugging Face transformer model

Initializing the T5-large transformer model

Getting started with T5

Exploring the architecture of the T5 model

Summarizing documents with T5-large

Creating a summarization function

A general topic sample

The Bill of Rights sample

A corporate law sample

From text-to-text to new word predictions with OpenAI ChatGPT

Comparing T5 and ChatGPT's summarization methods

Pretraining

Specific versus non-specific tasks

Summarization with ChatGPT

Summary

Questions

References

Further reading

14. Exploring Cutting-Edge LLMs with Vertex AI and PaLM 2

Architecture

Pathways

Client

Resource manager

Intermediate representation

Compiler

Scheduler

Executor

PaLM

Parallel layer processing that increases training speed

Shared input-output embeddings, which saves memory
No biases, which improves training stability
Rotary Positional Embedding (RoPE) improves model quality
SwiGLU activations improve model quality

PaLM 2

Improved performance, faster, and more efficient
Scaling laws, optimal model size, and the number of parameters
State-of-the-art (SOA) performance and a new training methodology

Assistants

Gemini
Google Workspace
Google Colab Copilot
Vertex AI PaLM 2 interface
Vertex AI PaLM 2 assistant

Vertex AI PaLM 2 API

Question answering
Question-answer task
Summarization of a conversation
Sentiment analysis
Multi-choice problems
Code

Fine-tuning

Creating a bucket
Fine-tuning the model

Summary

Questions

References

Further reading

15. Guarding the Giants: Mitigating Risks in Large Language Models
 - The emergence of functional AGI
 - Cutting-edge platform installation limitations
 - Auto-BIG-bench
 - WandB
 - When will AI agents replicate?

Function: `create_vocab`

Process:

Function: `scrape_wikipedia`

Process:

Function: `create_dataset`

Process:

Classes: `TextDataset`, `Encoder`, and `Decoder`

Function: `count_parameters`

Function: `main`

Process:

Saving and Executing the Model

Risk management

Hallucinations and memorization

Memorization

Risky emergent behaviors

Disinformation

Influence operations

Harmful content

Privacy

Cybersecurity

Risk mitigation tools with RLHF and RAG

1. Input and output moderation with transformers and a rule base

2. Building a knowledge base for ChatGPT and GPT-4
 Adding keywords

3. Parsing the user requests and accessing the KB

4. Generating ChatGPT content with a dialog function

 Token control

 Moderation

Summary

Questions

References

Further reading

16. Beyond Text: Vision Transformers in the Dawn of Revolutionary AI

From task-agnostic models to multimodal vision transformers

ViT – Vision Transformer

The basic architecture of ViT

[Step 1: Splitting the image into patches](#)
[Step 2: Building a vocabulary of image patches](#)
[Step 3: The transformer](#)

[Vision transformers in code](#)
[A feature extractor simulator](#)
[The transformer](#)
[Configuration and shapes](#)

[CLIP](#)

[The basic architecture of CLIP](#)
[CLIP in code](#)

[DALL-E 2 and DALL-E 3](#)

[The basic architecture of DALL-E](#)
[Getting started with the DALL-E 2 and DALL-E 3 API](#)

[Creating a new image](#)

[Creating a variation of an image](#)

[From research to mainstream AI with DALL-E](#)

[GPT-4V, DALL-E 3, and divergent semantic association](#)

[Defining divergent semantic association](#)

[Creating an image with ChatGPT Plus with DALL-E](#)

[Implementing the GPT-4V API and experimenting with DAT](#)

[Example 1: A standard image and text](#)

[Example 2: Divergent semantic association, moderate divergence](#)

[Example 3: Divergent semantic association, high divergence](#)

[Summary](#)

[Questions](#)

[References](#)

[Further Reading](#)

[17. Transcending the Image-Text Boundary with Stable Diffusion](#)

[Transcending image generation boundaries](#)

[Part I: Defining text-to-image with Stable Diffusion](#)

- [1. Text embedding using a transformer encoder](#)
- [2. Random image creation with noise](#)
- [3. Stable Diffusion model downsampling](#)
- [4. Decoder upsampling](#)
- [5. Output image](#)

[Running the Keras Stable Diffusion implementation](#)

[Part II: Running text-to-image with Stable Diffusion](#)

[Generative AI Stable Diffusion for a Divergent Association Task \(DAT\)](#)

[Part III: Video](#)

[Text-to-video with Stability AI animation](#)

[Text-to-video, with a variation of OpenAI CLIP](#)

[A video-to-text model with TimeSformer](#)

[Preparing the video frames](#)

[Putting the TimeSformer to work to make predictions on the video frames](#)

[Summary](#)

[Questions](#)

[References](#)

[Further reading](#)

18. [Hugging Face AutoTrain: Training Vision Models without Coding](#)

[Goal and scope of this chapter](#)

[Getting started](#)

[Uploading the dataset](#)

[No coding?](#)

[Training models with AutoTrain](#)

[Deploying a model](#)

[Running our models for inference](#)

[Retrieving validation images](#)

[The program will now attempt to classify the validation images. We will see how a vision transformer reacts to this image.](#)

[Inference: image classification](#)

[Validation experimentation on the trained models](#)

[ViTForImageClassification](#)

[SwinForImageClassification 1](#)

[BeitForImage Classification](#)

[SwinForImageClassification 2](#)

[ConvNextForImageClassification](#)

[ResNetForImageClassification](#)

[Trying the top ViT model with a corpus](#)

[Summary](#)

[Questions](#)

[References](#)

[Further reading](#)

19. [On the Road to Functional AGI with HuggingGPT and its Peers](#)

[Defining F-AGI](#)

[Installing and importing](#)

[Validation set](#)

[Level 1 image: easy](#)

[Level 2 image: difficult](#)

[Level 3 image: very difficult](#)

[HuggingGPT](#)

[Level 1: Easy](#)

[Level 2: Difficult](#)

[Level 3: Very difficult](#)

[CustomGPT](#)

[Google Cloud Vision](#)

[Level 1: Easy](#)

[Level 2: Difficult](#)

[Level 3: Very difficult](#)

[Model chaining: Chaining Google Cloud Vision to ChatGPT](#)

[Model Chaining with Runway Gen-2](#)

[Midjourney: Imagine a ship in the galaxy](#)

[Gen-2: Make this ship sail the sea](#)

[Summary](#)

[Questions](#)

[References](#)

[Further Reading](#)

20. [Beyond Human-Designed Prompts with Generative Ideation](#)

[Part I: Defining generative ideation](#)

[Automated ideation architecture](#)

[Scope and limitations](#)

[Part II: Automating prompt design for generative image design](#)

[ChatGPT/GPT-4 HTML presentation](#)

[ChatGPT with GPT-4 provides the text for the presentation](#)

[ChatGPT with GPT-4 provides a graph in HTML to illustrate the presentation](#)

[Llama 2](#)

[A brief introduction to Llama 2](#)

[Implementing Llama 2 with Hugging Face Midjourney](#)

[Discord API for Midjourney](#)

[Microsoft Designer](#)

[Part III: Automated generative ideation with Stable Diffusion](#)

[1. No prompt: Automated instruction for GPT-4](#)

[2. Generative AI \(prompt generation\) using ChatGPT with GPT-4](#)

[3. and 4. Generative AI with Stable Diffusion and displaying images](#)

[The future is yours!](#)

[The future of development through VR-AI](#)

[The groundbreaking shift: Parallelization of development through the fusion of VR and AI](#)

[Opportunities and risks](#)

[Summary](#)

[Questions](#)

[References](#)

[Further reading](#)

[Appendix: Answers to the Questions](#)

[Other Books You May Enjoy](#)

[Index](#)

Preface

Transformer-driven Generative AI models are a game-changer for **Natural Language Processing (NLP)** and computer vision. Large Language Generative AI transformer models have achieved superhuman performance through services such as ChatGPT with GPT-4V for text, image, data science, and hundreds of domains. We have gone from primitive Generative AI to superhuman AI performance in just a few years!

Language understanding has become the pillar of language modeling, chatbots, personal assistants, question answering, text summarizing, speech-to-text, sentiment analysis, machine translation, and more. The expansion from the early **Large Language Models (LLMs)** to multimodal (text, image, sound) algorithms has taken AI into a new era.

For the past few years, we have been witnessing the expansion of social networks versus physical encounters, e-commerce versus physical shopping, digital newspapers, streaming versus physical theaters, remote doctor consultations versus physical visits, remote work instead of on-site tasks, and similar trends in hundreds more domains. This digital activity is now increasingly driven by transformer copilots in hundreds of applications.

The transformer architecture began just a few years ago as revolutionary and disruptive. It broke with the past, leaving the dominance of RNNs and CNNs behind. BERT and GPT models abandoned recurrent network layers and replaced them with self-attention. But in 2023, OpenAI GPT-4

proposed AI into new realms with GPT-4V (vision transformer), which is paving the path for functional (everyday tasks) AGI. Google Vertex AI offered similar technology. 2024 is not a new year in AI; it's a new decade! Meta (formerly Facebook) has released Llama 2, which we can deploy seamlessly on Hugging Face.

Transformer encoders and decoders contain attention heads that train separately, parallelizing cutting-edge hardware. Attention heads can run on separate GPUs, opening the door to billion-parameter models and soon-to-come trillion-parameter models.

The increasing amount of data requires training AI models at scale. As such, transformers pave the way to a new era of parameter-driven AI. Learning to understand how hundreds of millions of words and images fit together requires a tremendous amount of parameters. Transformer models such as Google Vertex AI PaLM 2 and OpenAI GPT-4V have taken emergence to another level. Transformers can perform hundreds of NLP tasks they were not trained for.

Transformers can also learn image classification and reconstruction by embedding images as sequences of words. This book will introduce you to cutting-edge computer vision transformers such as **Vision Transformers (ViTs)**, CLIP, GPT-4V, DALL-E 3, and Stable Diffusion.

Think of how many humans it would take to control the content of the billions of messages posted on social networks per day to decide if they are legal and ethical before extracting the information they contain.

Think of how many humans would be required to translate the millions of pages published each day on the web. Or imagine how many people it would take to manually control the millions of messages and images made per minute!

Imagine how many humans it would take to write the transcripts of all of the vast amount of hours of streaming published per day on the web.

Finally, think about the human resources that would be required to replace AI image captioning for the billions of images that continuously appear online.

This book will take you from developing code to prompt engineering, a new “programming” skill that controls the behavior of a transformer model.

Each chapter will take you through the key aspects of language understanding and computer vision from scratch in Python, PyTorch, and TensorFlow.

You will learn the architecture of the Original Transformer, Google BERT, GPT-4, PaLM 2, T5, ViT, Stable Diffusion, and several other models. You will fine-tune transformers, train models from scratch, and learn to use powerful APIs.

You will keep close to the market and its demand for language understanding in many fields, such as media, social media, and research papers, for example. You will learn how to improve Generative AI models with **Retrieval Augmented Generation (RAG)**, embedding-based searches, prompt engineering, and automated ideation with AI-generated prompts.

Throughout the book, you will work hands-on with Python, PyTorch, and TensorFlow. You will be introduced to the key AI language understanding neural network models. You will then learn how to explore and implement transformers.

You will learn the skills required not only to adapt to the present market but also to acquire the vision to face innovative projects and AI evolutions. This

book aims to give readers both the knowledge and the vision to select the right models and environment for any given project.

Who this book is for

This book is not an introduction to Python programming or machine learning concepts. Instead, it focuses on deep learning for machine translation, speech-to-text, text-to-speech, language modeling, question answering, and many more NLP domains, as well as computer vision multimodal tasks.

Readers who can benefit the most from this book are:

- Deep learning, vision, and NLP practitioners familiar with Python programming.
- Data analysts, data scientists, and machine learning/AI engineers who want to understand how to process and interrogate the increasing amounts of language-driven and image data.

What this book covers

Part I: The Foundations of Transformers

Chapter 1, What Are Transformers?, explains, at a high level, what transformers and Foundation Models are. We will first unveil the incredible power of the deceptively simple O(1) time complexity of transformer models that changed everything. We will continue to discover how a hardly known transformer algorithm in 2017 rose to dominate so many domains and brought us Foundation Models.

Chapter 2, Getting Started with the Architecture of the Transformer Model, goes through the background of NLP to understand how RNN, LSTM, and CNN architectures were abandoned and how the transformer architecture opened a new era. We will go through the Original Transformer's architecture through the unique *Attention Is All You Need* approach invented by the Google Research and Google Brain authors. We will describe the theory of transformers. We will get our hands dirty in Python to see how multi-attention head sublayers work.

Chapter 3, Emergent vs Downstream Tasks: The Unseen Depths of Transformers, bridges the gap between the functional and mathematical architecture of transformers by introducing *emergence*. We will then see how to measure the performance of transformers before exploring several downstream tasks, such as the **Standard Sentiment TreeBank (SST-2)**, linguistic acceptability, and Winograd schemas.

Chapter 4, Advancements in Translations with Google Trax, Google Translate, and Gemini, goes through machine translation in three steps. We will first define what machine translation is. We will then preprocess a

Workshop on Machine Translation (WMT) dataset. Finally, we will see how to implement machine translations.

Chapter 5, Diving into Fine-Tuning through BERT, builds on the architecture of the Original Transformer. **Bidirectional Encoder Representations from Transformers (BERT)** takes transformers into a vast new way of perceiving the world of NLP. Instead of analyzing a past sequence to predict a future sequence, BERT attends to the whole sequence! We will first go through the key innovations of BERT's architecture and then fine-tune a BERT model by going through each step in a Google Colaboratory notebook. Like humans, BERT can learn tasks and perform other new ones without having to learn the topic from scratch.

Chapter 6, Pretraining a Transformer from Scratch through RoBERTa, builds a RoBERTa transformer model from scratch using the Hugging Face PyTorch modules. The transformer will be both BERT-like and DistilBERT-like. First, we will train a tokenizer from scratch on a customized dataset. Finally, we will put the knowledge acquired in this chapter to work and pretrain a Generative AI customer support model on X (formerly Twitter) data.

Part II: The Rise of Suprahuman NLP

Chapter 7, The Generative AI Revolution with ChatGPT, goes through the tremendous improvements and diffusion of ChatGPT models into the everyday lives of developers and end-users. We will first examine the architecture of OpenAI's GPT models before working with the GPT-4 API and its hyperparameters to implement several NLP examples. Finally, we will learn how to obtain better results with **Retrieval Augmented Generation (RAG)**. We will implement an example of automated RAG with GPT-4.

Chapter 8, Fine-Tuning OpenAI GPT Models, explores fine-tuning to make sense of the choices we can make for a project to go in this direction or not. We will introduce risk management perspectives. We will prepare a dataset and fine-tune a cost-effective babbage-02 model for a completion task.

Chapter 9, Shattering the Black Box with Interpretable Tools, lifts the lid on the black box that is transformer models by visualizing their activity. We will use BertViz to visualize attention heads, **Language Interpretability Tool (LIT)** to carry out a **Principal Component Analysis (PCA)**, and LIME to visualize transformers via dictionary learning. OpenAI LLMs will take us deeper and visualize the activity of a neuron in a transformer with an interactive interface. This approach opens the door to GPT-4 explaining a transformer, for example.

Chapter 10, Investigating the Role of Tokenizers in Shaping Transformer Models, introduces some tokenizer-agnostic best practices to measure the quality of a tokenizer. We will describe basic guidelines for datasets and tokenizers from a tokenization perspective. We will explore word and subword tokenizers and show how a tokenizer can shape a transformer

model's training and performance. Finally, we will build a function to display and control token-ID mappings.

Chapter 11, Leveraging LLM Embeddings as an Alternative to Fine-Tuning, explains why searching with embeddings can sometimes be a very effective alternative to fine-tuning. We will go through the advantages and limits of this approach. We will go through the fundamentals of text embeddings. We will build a program that reads a file, tokenizes it, and embeds it with Gensim and Word2Vec. We will implement a question-answering program on sports events and use OpenAI Ada to embed Amazon fine food reviews. By the end of the chapter, we will have taken a system from prompt design to advanced prompt engineering using embeddings for RAG.

Chapter 12, Toward Syntax-Free Semantic Role Labeling with ChatGPT and GPT-4, goes through the revolutionary concepts of syntax-free, nonrepetitive stochastic models. We will use ChatGPT Plus with GPT-4 to run easy to complex **Semantic Role Labeling (SRL)** samples. We will see how a general-purpose, emergent model reacts to our SRL requests. We will progressively push the transformer model to the limits of SRL.

Chapter 13, Summarization with T5 and ChatGPT, goes through the concepts and architecture of the T5 transformer model. We will then apply T5 to summarize documents with Hugging Face models. The examples in this chapter will be legal and medical to explore domain-specific summarization beyond simple texts. We are not looking for an easy way to implement NLP but preparing ourselves for the reality of real-life projects. We will then compare T5 and ChatGPT approaches to summarization.

Chapter 14, Exploring Cutting-Edge LLMs with Vertex AI and PaLM 2, examines Pathways to understand PaLM. We will continue and look at the main features of **PaLM (Pathways Language Model)**, a decoder-only,

densely activated, and autoregressive transformer model with 540 billion parameters trained on Google’s Pathways system. We will see how Google PaLM 2 can perform a chat task, a discriminative task (such as classification), a completion task (also known as a generative task), and more. We will implement the Vertex AI PaLM 2 API for several NLP tasks, including question-answering and summarization. Finally, we will go through Google Cloud’s fine-tuning process.

Chapter 15, Guarding the Giants: Mitigating Risks in Large Language Models, examines the risks of LLMs, risk management, and risk mitigation tools. The chapter explains hallucinations, memorization, risky emergent behavior, disinformation, influence operations, harmful content, adversarial attacks (“jailbreaks”), privacy, cybersecurity, overreliance, and memorization. We will then go through some risk mitigation tools through advanced prompt engineering, such as implementing a moderation model, a knowledge base, keyword parsing, prompt pilots, post-processing moderation, and embeddings.

Part III: Generative Computer Vision: A New Way to See the World

Chapter 16, Beyond Text: Vision Transformers in the Dawn of Revolutionary AI, explores the innovative transformer models that respect the basic structure of the Original Transformer but make some significant changes. We will discover powerful computer vision transformers like ViT, CLIP, DALL-E, and GPT-4V. We will implement vision transformers in code, including GPT-4V, and expand the text-image interactions of DALL-3 to divergent semantic association. We will take OpenAI models into the nascent world of highly divergent semantic association creativity.

Chapter 17, Transcending the Image-Text Boundary with Stable Diffusion, delves into diffusion models, introducing Stable Vision, which has created a disruptive generative image AI wave rippling through the market. We will then dive into the principles, math, and code of the remarkable Keras Stable Diffusion model. We will go through each of the main components of a Stable Diffusion model and peek into the source code provided by Keras and run the model. We will run a text-to-video synthesis model with Hugging Face and a video-to-text task with Meta's TimeSformer.

Chapter 18, Hugging Face AutoTrain: Training Vision Models without Coding, explores how to train a vision transformer using Hugging Face's AutoTrain. We will go through the automated training process and discover the unpredictable problems that show why even automated ML requires human AI expertise. The goal of this chapter is also to show how to probe the limits of a computer vision model, no matter how sophisticated it is.

Chapter 19, On the Road to Functional AGI with HuggingGPT and its Peers, shows how we can use cross-platform chained models to solve difficult image classification problems. We will put HuggingGPT and Google Cloud Vision to work to identify easy, difficult, and very difficult images. We will go beyond classical pipelines and explore how to chain heterogeneous competing models.

Chapter 20, Beyond Human-Designed Prompts with Generative Ideation, explores generative ideation, an ecosystem that automates the production of an idea to text and image content. The development phase requires highly skilled human AI experts. For an end user, the ecosystem is a click-and-run experience. By the end of this chapter, we will be able to deliver ethical, exciting, generative ideation to companies with no marketing resources. We will be able to expand generative ideation to any field in an exciting, cutting-edge, yet ethical ecosystem.

To get the most out of this book

Most of the programs in the book are Jupyter notebooks. All you will need is a free Google Gmail account, and you will be able to run the notebooks on Google Colaboratory's free VM.

Take the time to read *Chapter 2, Getting Started with the Architecture of the Transformer Model*. Chapter 2 contains the description of the Original Transformer. If you find it difficult, then pick up the general intuitive ideas from the chapter. You can then go back to that chapter when you feel more comfortable with transformers after a few chapters.

After reading each chapter, consider how you could implement transformers for your customers or use them to move up in your career with novel ideas.

Download the example code files

The code bundle for the book is hosted on GitHub at

<https://github.com/Denis2054/Transformers-for-NLP-and-Computer-Vision-3rd-Edition>. We also have other code bundles from our rich catalog of books and videos available at
<https://github.com/PacktPublishing/>. Check them out!

Download the color images

We also provide a PDF file that contains color images of the screenshots/diagrams used in this book. You can download it here:
<https://packt.link/gbp/9781805128724>.

Conventions used

There are several text conventions used throughout this book.

`CodeInText`: Indicates sentences and words run through the models in the book, code words in text, database table names, folder names, filenames, file extensions, pathnames, dummy URLs, user input, and Twitter handles. For example, “However, if you wish to explore the code, you will find it in the Google Colaboratory `positional_encoding.ipynb` notebook and the `text.txt` file in this chapter’s GitHub repository.”

A block of code is set as follows:

```
import numpy as np
from scipy.special import softmax
```

When we wish to draw your attention to a particular part of a code block, the relevant lines or items are set in bold:

```
The black cat sat on the couch and the brown dog slept on the
```

Any command-line input or output is written as follows:

```
vector similarity  
[[0.9627094]] final positional encoding similarity
```

Bold: Indicates a new term, an important word, or words that you see on the screen.

For instance, words in menus or dialog boxes also appear in the text like this. For example:

“In our case, we are looking for **t5-large**, a t5-large model we can smoothly run in

Google Colaboratory.”



Warnings or important notes appear like this.



Tips and tricks appear like this.

Get in touch

Feedback from our readers is always welcome.

General feedback: Email feedback@packtpub.com and mention the book's title in the subject of your message. If you have questions about any aspect of this book, please email us at questions@packtpub.com.

Errata: Although we have taken every care to ensure the accuracy of our content, mistakes do happen. If you have found a mistake in this book, we would be grateful if you would report this to us. Please visit <http://www.packtpub.com/submit-errata>, selecting your book, clicking on the Errata Submission Form link, and entering the details.

Piracy: If you come across any illegal copies of our works in any form on the internet, we would be grateful if you would provide us with the location address or website name. Please contact us at copyright@packtpub.com with a link to the material.

If you are interested in becoming an author: If there is a topic that you have expertise in and you are interested in either writing or contributing to a book, please visit <http://authors.packtpub.com>.

Share your thoughts

Once you've read *Transformers for Natural Language Processing and Computer Vision - Third Edition*, we'd love to hear your thoughts! Please [click here to go straight to the Amazon review](#) page for this book and share your feedback.

Your review is important to us and the tech community and will help us make sure we're delivering excellent quality content.

Download a free PDF copy of this book

Thanks for purchasing this book!

Do you like to read on the go but are unable to carry your print books everywhere?

Is your eBook purchase not compatible with the device of your choice?

Don't worry, now with every Packt book you get a DRM-free PDF version of that book at no cost.

Read anywhere, any place, on any device. Search, copy, and paste code from your favorite technical books directly into your application.

The perks don't stop there, you can get exclusive access to discounts, newsletters, and great free content in your inbox daily

Follow these simple steps to get the benefits:

1. Scan the QR code or visit the link below



<https://packt.link/free-ebook/9781805128724>

2. Submit your proof of purchase
3. That's it! We'll send your free PDF and other benefits to your email directly

1

What Are Transformers?

Transformers are industrialized, homogenized **Large Language Models (LLMs)** designed for parallel computing. A transformer model can carry out a wide range of tasks with no fine-tuning. Transformers can perform self-supervised learning on billions of records of raw unlabeled data with billions of parameters. From these billion-parameter models emerged multimodal architectures that can process text, images, audio, and videos.

ChatGPT popularized the usage of transformer architectures that have become general-purpose technologies like printing, electricity, and computers.

Applications are burgeoning everywhere! Google Cloud AI, **Amazon Web Services (AWS)**, Microsoft Azure, OpenAI, Google Workspace, Microsoft 365, Google Colab Copilot, GitHub Copilot, Hugging Face, Meta, and myriad other offers are emerging.

The functionality of transformer models has pervaded every aspect of our workspaces with Generative AI for text, Generative AI for images, discriminative AI, task specific-models, unsupervised learning, supervised learning, prompt design, prompt engineering, text-to-code, code-to-text, and more. Sometimes, a GPT-like model will encompass all these concepts!

The societal impact is tremendous. Developing an application has become an educational exercise in many cases. A project manager can now go to

OpenAI's cloud platform, sign up, obtain an API key, and get to work in a few minutes. Users can then enter a text, specify the NLP task as Google Workspace or Microsoft 365, and obtain a response created by a Google Vertex AI or a ChatGPT transformer model. Finally, users can go to Google's Gen App Builder and build applications without programming or machine learning knowledge.

The numbers are dizzying. *Bommasani et al.* (2023) created a Foundation Model ecosystem that lists 128 Foundation Models 70 applications, and 64 datasets. The paper also mentions Hugging Face's 150,000+ models and 20,000+ datasets! The list is growing weekly and will spread to every activity in society.

Where does that leave an AI professional or someone wanting to be one?

Should a project manager choose to work locally? Or should the implementation be done directly on Google Cloud, Microsoft Azure, or AWS? Should a development team select Hugging Face, Google Trax, OpenAI, or AllenNLP? Should an AI professional use an API with practically no AI development? Should an end-user build a no-code AI application with no ML knowledge with Google's Gen App Builder?

The answer is yes to *all* of the above! You do not know what a future employer, customer, or user may want or specify. Therefore, you must be ready to adapt to any need that comes up at the dataset, model, and application levels. This book does not describe all the offers that exist on the market. You cannot learn every single model and platform on the market. If you try to learn everything, you'll remember nothing. You need to know where to start and when to stop. By the book's end, you will have acquired enough critical knowledge to adapt to this ever-moving market.

In this chapter, we will first unveil the incredible power of the deceptively simple $O(1)$ time complexity of transformer models that changed everything. We will build a notebook in Python, PyTorch, and TensorFlow to see how transformers hijacked hardware accelerators. We will then discover how one token (a minimal part of a word) led to the AI revolution we are experiencing.

We will continue to discover how a hardly known transformer algorithm in 2017 rose to dominate so many domains. We had to find a new name for it: the Foundation Model. Foundation Models can do nearly everything in AI! We will sit back and watch how ChatGPT explains, analyzes, writes a classification program in a Python notebook, and displays a decision tree.

Finally, this chapter introduces the role of an AI professional in the ever-changing job market. We will begin to tackle the problem of choosing the right resources.

We must address these critical notions before starting our exploratory journey through the variety of transformer model implementations described in this book.

This chapter covers the following topics:

- How one $O(1)$ invention changed the course of AI history
- How transformer models hijacked hardware accelerators
- How one token overthrew hundreds of AI applications
- The multiple facets of a transformer model
- Generative AI versus discriminative AI
- Unsupervised and self-supervised learning versus supervised learning
- General-purpose models versus task-specific models
- How ChatGPT has changed the meaning of automation

Another random document with
no related content on Scribd:

THE
AUSTRALIAN
ABORIGINAL

BY
HERBERT BASEDOW
M.A., M.D., Ph.D., B.Sc., F.G.S., ETC.

SOMETIMES CHIEF MEDICAL INSPECTOR AND CHIEF PROTECTOR OF ABORIGINES
IN THE NORTHERN TERRITORY; SPECIAL ABORIGINES' COMMISSIONER
FOR THE FEDERAL AND STATE GOVERNMENTS; LOCAL CORRE-
SPONDENT ROYAL ANTHROPOLOGICAL INSTITUTE OF GREAT
BRITAIN AND IRELAND; HONORARY FELLOW ANTHRO-
POLOGICAL SOCIETY OF GÖTTINGEN, ETC.

With 146 Illustrations

ADELAIDE:
F. W. PREECE AND SONS
1925

THE
AUSTRALIAN
ABORIGINAL

BY

HERBERT BASEDOW

M.A., M.D., Ph.D., B.Sc., F.G.S., etc.

SOMETIMES CHIEF MEDICAL INSPECTOR AND CHIEF PROTECTOR OF ABORIGINES
IN THE NORTHERN TERRITORY; SPECIAL ABORIGINES' COMMISSIONER
FOR THE FEDERAL AND STATE GOVERNMENTS; LOCAL CORRESPONDENT
ROYAL ANTHROPOLOGICAL INSTITUTE OF GREAT
BRITAIN AND IRELAND; HONORARY FELLOW ANTHROPOLOGICAL
SOCIETY OF GÖTTINGEN, ETC.

With 146 Illustrations

ADELAIDE:

F. W. PREECE AND SONS

1925

The Project Gutenberg eBook of The Australian aboriginal

This ebook is for the use of anyone anywhere in the United States and most other parts of the world at no cost and with almost no restrictions whatsoever. You may copy it, give it away or re-use it under the terms of the Project Gutenberg License included with this ebook or online at www.gutenberg.org. If you are not located in the United States, you will have to check the laws of the country where you are located before using this eBook.

Title: The Australian aboriginal

Author: Herbert Basedow

Release date: October 23, 2023 [eBook #71940]

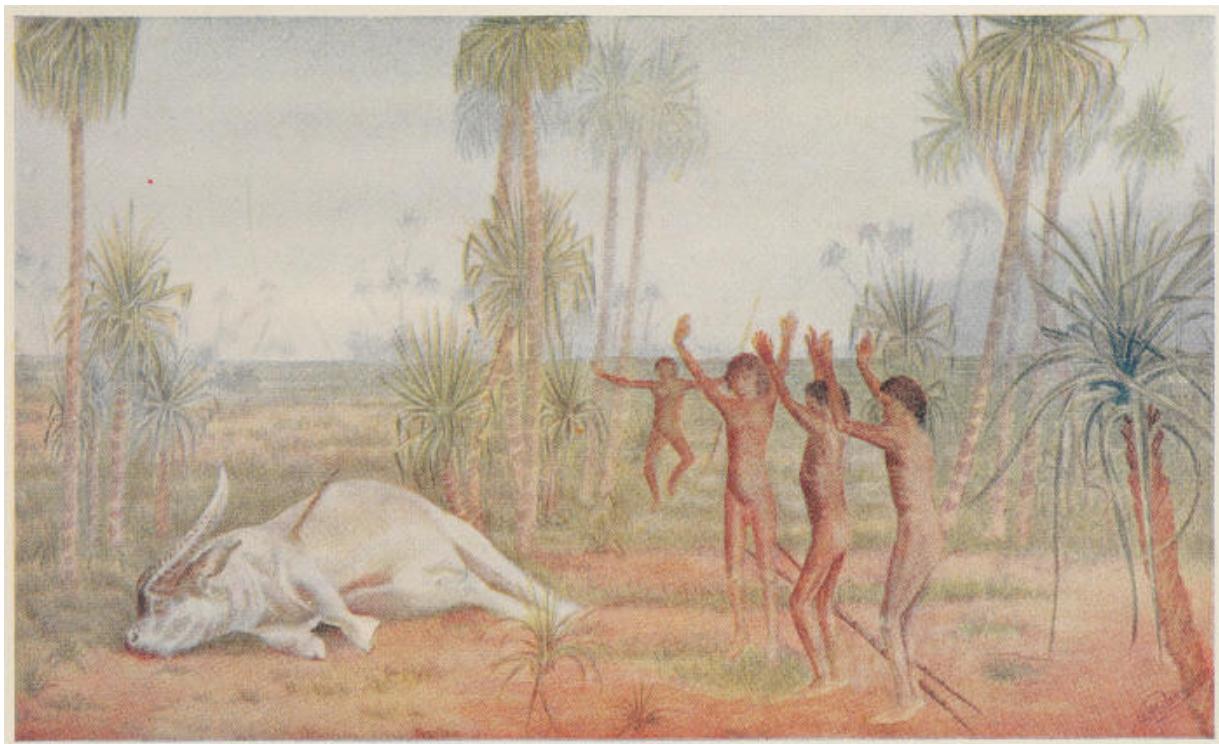
Language: English

Original publication: Adelaide: F. W. Freece & Sons, 1925

Credits: Tim Lindell and the Online Distributed Proofreading Team at <https://www.pgdp.net> (This file was produced from images generously made available by The Internet Archive)

*** START OF THE PROJECT GUTENBERG EBOOK THE
AUSTRALIAN ABORIGINAL ***

THE AUSTRALIAN ABORIGINAL



THE BUFFALO HUNTERS' TRIUMPH

H. BASEDOW, *pinx.*

THE
AUSTRALIAN
ABORIGINAL

BY
HERBERT BASEDOW
M.A., M.D., PH.D., B.Sc., F.G.S., ETC.
SOMETIME CHIEF MEDICAL INSPECTOR AND CHIEF PROTECTOR OF
ABORIGINES
IN THE NORTHERN TERRITORY; SPECIAL ABORIGINES' COMMISSIONER
FOR THE FEDERAL AND STATE GOVERNMENTS; LOCAL CORRESPONDENT
ROYAL ANTHROPOLOGICAL INSTITUTE OF GREAT
BRITAIN AND IRELAND; HONORARY FELLOW ANTHROPOLOGICAL
SOCIETY OF GÖTTINGEN, ETC.

With 146 Illustrations

ADELAIDE:
F. W. PREECE AND SONS
1925

Registered by the Postmaster-General for
transmission through the post as a book.

Wholly set up and printed in Australia
at THE HASSELL PRESS, Adelaide.

TO THE MEMORY OF MY MOTHER
ANNA CLARA HELENE BASEDOW

PREFACE

For some years past it had been my intention to write a progressive series of treatises on the Australian aboriginal, embodying observations as they were being made. As time wore on, however, my procrastination amassed so great a fund and so great a variety of notes that my original plan became impossible. The delay was not altogether through any fault of mine, but through a run of fortunate happenings which allowed me to spend the better part of every year in the great unbounded spaces of central and northern Australia, to live among the uncontaminated tribes, and to study Australian anthropology at the fountain head. It was through my repeated and prolonged absences from the city (and civilization in general) that I could not attend to the publication of any lengthy scientific papers.

Apart from my private and professional journeys, the following are some of the better known expeditions I have accompanied or led:

The Government North-West Expedition,

Government Expedition of Geological Exploration in the Northern Territory,

Cruise of the s.s. "White Star" under His Excellency Sir George Le Hunte, Governor of South Australia,

Mineralogical Survey of the Flinders Ranges,

Sir Joseph Verco's Deep Sea Dredging Expeditions,

Exploration of Melville and Bathurst Islands,

Expedition in Search of Munition Minerals in the Northern Kimberleys of Western Australia, officially subsidized by the Premier (Hon. John Scaddan, M.P.),

Expedition in Search of Munition Minerals in Central Australia,

Three South Australian Medical Relief Expeditions among the Aborigines,

Expedition in Search of Reported Remains of Leichhardt, under auspices of S.A. Geographical Society,

Commonwealth Medical Survey of Aborigines in the Northern Territory,

Mararoa Geological Expedition from Darwin to Northern Kimberleys,

Viceroyal Expedition to Central Australia, under His Excellency, Sir Tom Bridges, Governor of South Australia,

Commonwealth Railways Commissioner's Journey from Mildura to Port Augusta, under Hon. P. G. Stewart and N. G. Bell, Esq.,

Expedition through the Interior of Australia, under His Excellency, the Earl of Stradbroke, Governor of Victoria.

During the terms that I held official positions, firstly as a State Geologist, and secondly as Chief Medical Officer and Chief Protector of Aborigines for the Commonwealth Government in the Northern Territory, opportunities were afforded me of continuing my investigations among the indigenous population of Australia; especially fruitful were my researches when, as a Special Aborigines' Commissioner, I medically overhauled the tribes of South Australia and of the southern region of the Northern Territory.

In a more private capacity the southern districts and goldfields of Western Australia, the coastal and south-western districts of Queensland, and the north-western areas of New South Wales were traversed. Quite recently, too, a professional excursion to Java considerably enriched my knowledge of Melanesian ethnography and helped to explain the existence of several cults in the northern districts of Australia which border on the Indian Ocean.

On the other hand, not long before the war I continued study abroad and was privileged to be associated with the late Professor Hermann Klaatsch, under whose admirable guidance my researches were conducted in the Anatomical School of the University of Breslau. It will be apparent, therefore, why many of the results enumerated in the chapters dealing with the racial characteristics of the Australian aboriginal are based upon the doctrines of this eminent authority. In London Sir Arthur Keith courteously placed the whole of the valuable collections of Australian skulls and skeletons in the Hunterian Museum of the Royal College of Surgeons at my disposal; while Captain T. A. Joyce, in an equally generous way, facilitated my investigations in the anthropological galleries of the British Museum; I am taking advantage of this opportunity of expressing my sincere gratitude to these two gentlemen. I also desire to thank the Directors of the National and University Museums which I visited in France, Germany, Switzerland, and Holland, who so willingly responded to my enquiries and brought me into personal contact with any matters I was especially interested in.

While engaged upon one of my later commissions in central Australia, my duties took me to the Hermannsburg Mission Station on the Finke River, where for a fortnight I enjoyed the companionship of the Rev. Carl Strehlow and received every assistance in my researches among the western groups of the Arunndta tribe. Through the recent pathetic and heroic death of Strehlow, Science has lost an indefatigable and conscientious worker, and the aborigines a staunch and faithful friend.

In the present volume I have endeavoured to sift my subject matter in such a way as to keep the text in a suitable sequence and to make it of general interest. The principal difficulty has been to delete matter in order to keep down the bulk of the book. The latter remark applies equally well to the selection of illustrations; it was with a heavy heart that I found myself obliged to reduce the number of plates, all of which illustrated interesting points referred to in the text. Nevertheless, I feel that I am greatly indebted to the publisher, Mr. F. W. Preece, for allowing me to include so many more illustrations than he had originally counted upon.

Now that the manuscript is in the hands of the printer, I feel that there are several chapters I should like to have done more justice to; those, for instance, dealing with tribal organizations, initiation ceremonies, religious ideas, and art might have been considerably expanded if it had not been for the want of space. In those on religious ideas and art, I realize that I am launching in new directions so far as Australian anthropology is concerned, but I trust that the evidence which I have produced will be sufficient to prove that I did not arrive at the results too hastily, or, indeed, without deliberation and substantiation. Phallic worship had long been suspected in Australia, and some of our foremost authorities, such as the late R. Etheridge, Jun., were inclined to explain the occurrence of certain stones and other objects in the tribal areas of Australia (present or past) on that hypothesis. I came upon the evidence accidentally in 1916 in the form of a stone phallus erected in the ground at Success Strait in the far north-western corner of Australia. The stone was surrounded by a cleared cirque where much blood had been spilt at a recent ceremony. A few weeks later I for the first time witnessed an actual performance on the shores of Cambridge Gulf at which wooden phallus were produced. Since then I have been able to trace the existence of phallic worship of some form or other in several districts of central and northern Australia, an outline of which appears in a subsequent chapter. I trust that the facts I have been able to collect may help to illumine the somewhat doubtful question of how an aboriginal looks upon the process of procreation. Phallicism is closely related to such forms of Nature worship as are practised in order to make any species of animal or plant proliferate, or, for that matter, to bring down a shower of rain in times of drought. The phallus might gradually merge into a tjuringa. The painted "*Ngadanji*" and "*Ilbarinam*" tjuringas of the Arunndta tribe are regarded as images of the reproductive organ of a spirit which can generate life; in that sense they might be classed as phallus. I find that ancestor worship is generally indulged in; it is difficult at times to distinguish between an original spirit ancestor and a deity, but a Supreme Spirit or Deity is believed to exist and to rule over all creation.

With regard to totemism, I have shown the beliefs in a manner slightly different from those hitherto recorded. The mystical relationship between individual and object is traced to a mythical semi-human forerunner which was common to both and is now a spirit. The Arunndta call the spirit "*Knaninja*." The animal or plant relative of an individual is what has been commonly called the "totem" in Australian ethnology, while the symbolic representation of this object, which becomes the crest of the individual, corresponds to the "*kobong*" first described by Sir George Grey. Emblematic representations of both the "totemic" object and of the *Knaninja* are embodied in the tjuringa of the individual.

The essay on art, it will be observed, has been written on evolutional lines with respect to both technique and design. It will be understood that the material upon which the conclusions are based was collected in many parts of Australia and during many years of travel. The cults in question are in many cases distributed all over the continent, but occasionally are quite local. As an instance of the latter kind, I mention the famous drawings recorded by Sir George Grey, some of which I was fortunate in being able to locate and study on the Glenelg River in the far north-west. I might mention that, with very few exceptions, the designs appearing in the book as text figures are actual tracings reduced by photography to a size in keeping with the dimensions of the page. There is no doubt that primitive art in Australia is a fascinating study which has not received the attention it merits; and unfortunately it is rather late in the day to think of making a start. The system of conventionalism derived from the numerous pictographs and carvings is the basis of all characters and messages one finds on sticks, stones, and persons; it is the only key to an understanding of sacred tjuringa symbols. We have for too long looked upon aboriginal designs as meaningless, and upon aboriginal art production as being idle concoctions out of nothing which were invented just to make a thing "look pretty." This is anything but the true position. An aboriginal artist knows no such thing as a design without motive or origin; to him the shortest line or the smallest circle conveys a thought. In the chapter dealing with stone implements, I have, among other things, described a new type

of scraper which was used by the now extinct Adelaide tribe for trimming skins of animals.

The article on language is not intended to be at all comprehensive; my main object was to give a general idea of the construction, together with a few examples of the idiomatic uses, of the aboriginal tongue. I hope at a later date to be able to present a complete account of the Arunndta and Aluridja dialects, including the syntactical rules and grammatical forms.

I have to offer an apology to any authors who may claim priority to some of the facts which I mention in this book. I have written this account of the Australian aboriginal without attempting to consult previous literature, for the simple reason that, had I started looking up all necessary references, the volume might never have been completed. My time at headquarters has been so limited during the last fifteen years that, in the absence of a library near at hand, it was impossible for me to adopt any other method than to write up my observations at first-hand and run the risk of a certain amount of trespass. Nevertheless, I trust that the authors so affected will realize that there was no slight intended and will treat my transgression in the spirit of independent corroboration.

Our knowledge of Australian ethnology is so meagre that every man who has had first-hand experience among the tribes should consider it his bounden duty to place on record any facts he possesses, however trivial they may be. Every year the number of people who have seen the unsophisticated savage is dwindling. When I look back to the time of my first meeting with the tribes of central Australia, just twenty years ago, and compare the conditions of then and now, I shudder to think how quickly the romance of aboriginal affairs, together with all the scientific treasures it encompassed, has vanished, and is now irretrievably lost to the world. The rising generation will not have the advantage of men of even our time. Bones, stone artefacts, and wooden implements will remain in our museums for ever, but the habits, laws, beliefs, and legends are doomed to rapid extinction.

I do not claim to be an initiated member of any tribe. To be candid, I several times tried to qualify by impressive exhibitions of surgical skill and exaggerated munificence, but, although I gained the confidence and goodwill of the old men, I was informed that I could only be accepted provided I passed through the different grades of initiation and submitted to the attendant mutilations in the orthodox way. The medicine men, however, usually claimed me as a "*Kata*" or colleague, and allowed me to witness most of their rituals and sacred ceremonies, which they carefully explained to me. In this way I was able, also, to secure a very great series of photographs depicting intimate scenes from aboriginal life, many of which are unique. The only photographs illustrating this book which were not taken by me are those reproduced in Plates [XLVIII](#) and [XLIX](#); for these I am indebted to the late Mr. Nicholas Holtze.

I could not allow this opportunity to pass without making brief reference to the causes of the early extinction which is threatening these inoffensive, useful, and scientifically important people. We have only to cast our eyes in the direction of any wave of settlement to behold the disastrous effects our occupation of the land has had on the natives. Take, for instance, the Lake Eyre region, which embraces the Dieri, Yantowannta, Ngameni, and Yauroworka tribes. Official reports show that only forty years ago the population was so dense that the white settlers asked for greater police protection; the four tribes mentioned numbered many thousands. During a recent survey on behalf of the Government, I could barely muster three hundred wretched and decrepit souls in this region, who, literally speaking, were waiting for a lingering death to relieve them of their misery. We content ourselves by saying that civilization is the cause of the increased mortality, no doubt a plausible but very vague explanation. As a result of my investigations, I venture the opinion that the factor which has wrought the greatest havoc among the tribes is disease. The principal scourges are syphilis, pulmonary tuberculosis, and trachoma. Unless we realize the obligations which rest on our shoulders and give our natives a medical protection similar to the successful measures adopted by the United States, Canada, and New Zealand, they will continue to vanish and soon be classed as an extinct race.

H. BASEDOW.

Kent Town, South Australia, 2nd November, 1924.

CONTENTS

CHAPTER	PAGE
I INTRODUCTION TO AN AUSTRALIAN TRIBE	1
II RACIAL CHARACTERISTICS	5
III THE BREAST AND ABDOMEN	18
IV THE FACE AND ITS SKELETON	22
V THE MOUTH	31
VI THE SKULL AND BRAIN	34
VII COLOUR OF ABORIGINAL'S SKIN	40
VIII THE HAIR	46
IX LIKELY ORIGIN OF THE AUSTRALIAN ABORIGINAL	52
X AN ABORIGINAL'S BIRTH	61
XI CHILDHOOD	69
XII THE DAY'S MARCH	91
XIII CAMP LIFE	100
XIV HUNTING	120
XV VEGETABLE DIET	148
XVI BEVERAGES	153
XVII PITJURI	155
XVIII NAVIGATION	158
XIX DUELS	165
XX WARFARE	183
XXI SPEARS	190
XXII SPEAR-THROWERS	199
XXIII BURIAL AND MOURNING CUSTOMS	203
XXIV TRIBAL ORGANIZATIONS	216
XXV TRIBAL ADMINISTRATION	225
XXVI INITIATION	230
XXVII RELIGIOUS IDEAS	257
XXVIII ABORIGINAL ART	297

XXIX STONE IMPLEMENTS	359
XXX MUSIC AND DANCE	371
XXXI LANGUAGE	386

LIST OF PLATES

**FRONTISPIECE (IN COLOUR)—THE BUFFALO HUNTERS'
TRIUMPH**

PLATE	OPPOSITE PAGE
I Wordaman natives on the march	4
II Comparison of European with Aboriginal figure	12
III 1. Colossal brow-ridge, Arunndta man. 2. Supra-orbital prominence, deep notch at root of nose, prognathism, and female beard	16
IV 1. Aluridja woman. 2. Wongapitcha warrior, so- called Semitic type	32
V Wordaman man, profile and full-face	36
VI 1. Wongapitcha woman, wearing " <i>ungwaina</i> " (nose-stick) and fur-string bandeau. 2. Wongapitcha woman, wearing bloodwood seed pendants called " <i>dindula</i> "	44
VII 1. Old Yantowannta man, showing a strong growth of hair covering the body. 2. Old Yantowannta man, showing peculiar method of wearing the beard	48
VIII Old Kai-Kai, the leading medicine man of the western Arunndta	64
IX 1. Men of Kolaia tribe, Cambridge Gulf, wearing the hair tied at the back around a pad of emu feathers. 2. Wongapitcha men wearing ornamental wooden hair-pins	68
X 1. A juvenile "blonde," Aluridja tribe. 2. Ponga- Ponga gin carrying pet opossum on her head while on the march	76

XI	Rocking a child to sleep, Sunday Island	80
XII	Juvenile Types. 1. Full-face, female, Wongkanguru tribe. 2. Profile, female, Aluridja tribe	96
XIII	1. The game of "gorri," Humbert River, Northern Territory. 2. A "Kutturu" duel, Aluridja tribe	100
XIV	1. Arunndta boy practising with toy shield and boomerang. 2. Wordaman warrior, holding prevalent north-western type of spear- thrower and wearing pubic fur tassel	108
XV	1. Framework of hut in course of construction, Cooper's Creek, S.W. Queensland. 2. Hut decked with porcupine grass, Arltunga district	112
XVI	1. Wongapitcha women carrying dogs across their backs. 2. Kolaia man standing in the characteristic bird-like attitude, Cambridge Gulf	128
XVII	Female wood-carriers, Aluridja tribe	132
XVIII	Two handfuls of witchedy grubs	140
XIX	1. Aluridja tree-climber. 2. Wordaman tree- climber	144
XX	Kangaroo hunters, Aluridja tribe	160
XXI	1. Arunndta girl digging "Yelka." 2. Arunndta gin cleaning "Yelka" in bark pitchi	164
XXII	1. Sunday Islander making fire by the twirling process during a ceremonial. 2. "Kaloa" or mangrove raft, Worora tribe, Glenelg River district	172
XXIII	Aluridja men "pointing" the bone	176
XXIV	A "boned" man, Minning tribe	192
XXV	1. Dieri grave, Lake Eyre district. 2. Yantowannta grave, Innamincka district	196
XXVI	1. Aluridja widow. 2. Yantowannta widow	204
XXVII	Tooth-rapping ceremony, Wongapitcha tribe	208

XXVIII 1. Tooth-rapping ceremony. 2. Sunday Islander, who has had the two upper medium incisors removed during his initiation	224
XXIX 1. Old men introducing a dance during an initiation ceremony, Kukata tribe. 2. Circumcision ceremony, Kukata tribe	228
XXX Circumcision of a Wogait boy	236
XXXI Melville Islander, full-face and profile	240
XXXII An episode of the great fire ceremony, Kolaia tribe	256
XXXIII Ceremonial venesection, Arunndta tribe. 1. The median basilic vein is being slit. 2. The blood which is spurting from the incision is being collected on a shield	260
XXXIV The "Tjilbakuta" of the great emu ceremony, Arunndta tribe	268
XXXV Flashlight photograph of "Illiya Tjuringa" or great emu ceremony, Arunndta tribe	272
XXXVI 1. An ordinary performer in the Ladjia or yam ceremony, wearing the "tdela" head-gear. 2. The impersonator of the "Kuta Knaninja" in the Ladjia or yam ceremony	288
XXXVII The sacred "Etominja," Arunndta tribe	292
XXXVIII 1. Singing to the presiding spirit or Knaninja of the old women or "Arrekutja Tjuringa." 2. Ceremonial head-gear ("Tjilba Purra") of phallic significance	300
XXXIX A disenchanted area, Victoria River district	304
XL 1. Rock-carving of human form, Port Hedland. 2. Rock-carvings of lizard, pubic tassel, and owl, Flinders Ranges	320
XLI 1. Rock-carvings (including platypus design), Flinders Ranges. 2. Rock-carvings, Flinders Ranges	324
XLII 1. Rock-carvings, Flinders Ranges. 2. Emu design carved into the butt of a boabab	332