

天下文化

遠見

黃仁勳傳

輝達創辦人如何打造
全球最搶手的晶片

The Thinking Machine

Jensen Huang, Nvidia,
and the World's Most Coveted Microchip

史帝芬·維特 Stephen Witt 著
廖月娟、林俊宏 譯



黃仁勳傳

輝達創辦人如何打造
全球最搶手的晶片

The Thinking Machine

Jensen Huang, Nvidia,
and the World's Most Coveted Microchip

史帝芬·維特 Stephen Witt 著
廖月娟、林俊宏 譯

AI

1963 2

1993

1990

2013

AI

2000

2012

neural network

AI

2024

AI

		A/B	2015
2024		93	
	1993		1993
			Lori
	Major League Baseball	MLB	San Francisco
Giants			

OceanofPDF.com

Marcus Aurelius

OceanofPDF.com

S&P 500

3

2023

AI

first principle

500

Denny's

1993

Nvidia

1990
parallel computing

Quake

2012

neural network

200

OpenAI

AI

Midjourney AI ChatGPT AI Copilot AI

AI AI

Samuel Brannan	1849	3
----------------	------	---

AI

IBM 1960 AI

Podcast

AI

AI

AI

AI

2023 5 AI

AI

1,000

The New Yorker

AI

AI

AI

OpenAI

AI

AI

1973

[OceanofPDF.com](https://oceanofpdf.com)

OceanofPDF.com

1

1973

Oneida Elementary

Oneida Baptist Institute

Ben

Bays

Clay County



1963 2

1945

1949

1960

Carrier

1

1973

Tacoma

James Anderson

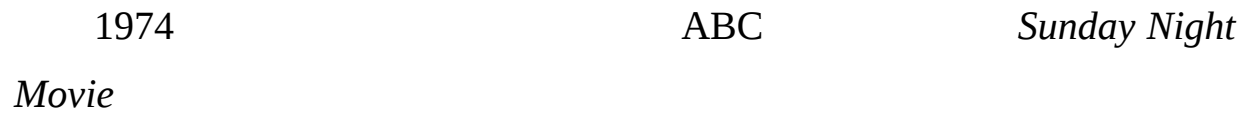
Burns 1899

1970

Jeff

Jensen

Jim



1 3,000

2020



1977	Apple II
<i>Super Star Trek</i>	Basic

Paddle Palace
Judy 1971

Lou Bochenski

Elks Lodge

2

Sports Illustrated

Joe

Romanosky

1970

3

University

Oregon State
Dean Verheiden

70

1980

Lori Mills

1980

1980

CD

1

2

3 National Honor Society 1921

OceanofPDF.com

1984

Supra

Advanced

Micro Devices AMD

AMD

28,700

AMD

AMD

Moscone Convention

Center

Chicago

Saturday in the Park

25 or 6 to 4

1984

AMD

DuPont

General Electric

James Stewart

Supra

Supra

Supra



AMD

AMD

AMD

AMD

1985

AMD

LSI Logic

1980

Very Large Scale Integration VLSI

1985

Silicon Graphics

SGI

3D

SGI

SGI AMD

101

San Jose

Palo Alto

40

Cupertino

Santa Clara

Milpitas

Mountain View

Intel

Cisco

8

1980

CRT

¹ LSI

1985

Jens

Horstmann

JH

Simulation Program with Integrated Circuit

Emphasis SPICE

SPICE

1 1 3

1 1 2 1.5

1988

Sushi

.....

1980

101 1992

Chris Malachowsky

Curtis Priem

Sun Microsystems

Rensselaer Polytechnic Institute

1970

1989 3D Sun GX

GX

1989

CRT

Sun GX

Wilf Corrigan

Sun GX

system-on-a-chip

3D

GX

1992

1990

Tina



2020

Hans Mosesmann

2.5

1 action office Robert Propst 1960 Herman Miller

[OceanofPDF.com](https://www.oceanofpdf.com)

3

1993

3.99

1993

3D

CGI

Jurassic Park

3D

graphics accelerator

Jon Peddie

5,000

5,000

5,000

Jim Gaither

NV new venture

NV1 envy Sun GX

NV Nvision

Nvidia invidia

Taco Bell

Berryessa

Road 2484

200

1993

2



GX

NV1

NV1

NV1

quadratic texture mapping

NV1

NV1

Sequoia Capital

Sutter Hill Ventures

Don Valentine

Mark Stevens

Tench Coxe

Patagonia



1994

Sunnyvale

NV1

Wells Fargo

Sega Sega

NV1

tape-out

Road Rash



Diamond

Jeff Fisher

VHS NV1

1990

Dwight Diercks

NV1 NV2 Sega
Dreamcast NV1 Sega
VR *Virtua Fighter* 3D
1995 3D 1993 9
Myst
Doom
NV1 1995
Matrox S3 3dfx Cirrus Logic ATI
Diamond Edge Diamond Monster 3dfx
NV1
Diamond Edge 249
Sega
II *Civilization II* *Command & Conquer*
NV1 VR
1996 VR
NV1 3D NV1
Tim Little depth buffer
NV1 Windows

NV1
Electronics

Diamond Edge

NV1

Fry's

1995

DirectX

NV1

NV1

NV1

Sega

NV2

NV3

OceanofPDF.com

4

David Kirk 1996

NV1

NV1

Sega

NV1

NV2

NV2

100

1996

Sega

Dreamcast

Sega

Sega

Cop

Sega

VR

Virtua



1997 NV3

NV3

NV3

NV3

128

NV3

DVD

NV3

Real-Time Interactive Video and Animation accelerator Riva 128

Best Buy 1997 8

Riva

Riva

NV1

NV3

Riva

5

1998

NV1

John

Sculley

Steve Jobs

1999

1998



T

Riva 128

1993

1998

Fountainhead

Douglas Adams

Isaac Asimov

The Hitchhiker's Guide to the Galaxy

Paulo Coelho

Ayn Rand

Foundation

The

Innovator's Dilemma

1997

Clayton Christensen

disruptive innovation

The

1960

The Beach Boys

General

Motors

Brian Wilson

2023



Riva

Taiwan Semiconductor Manufacturing Corporation TSMC

1931

Texas Instruments

1970

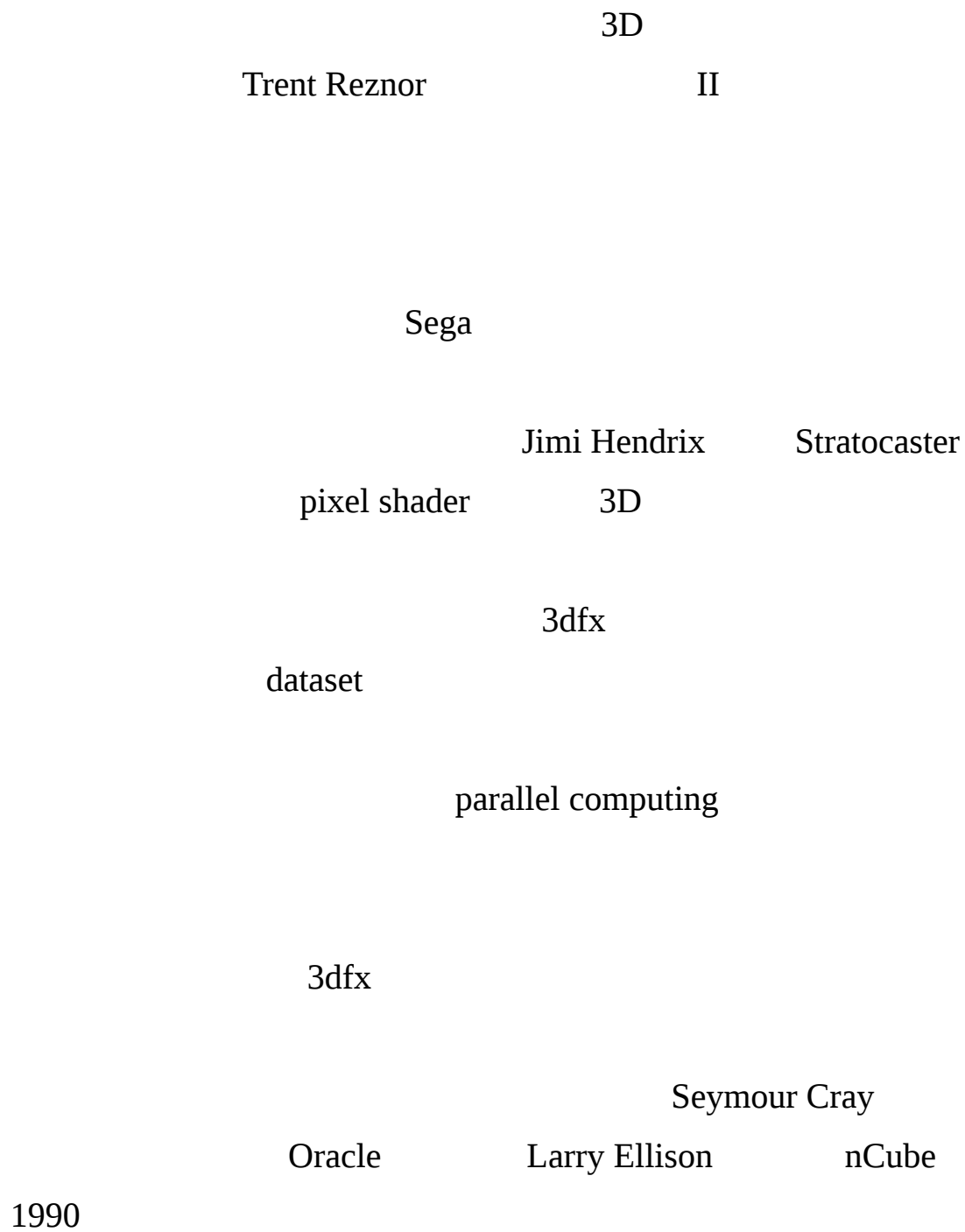
2010 0.3 5 Jane Hyun
Bamboo Ceiling

1998

1990



John Carmack *Doom*



TNT
TNT

OceanofPDF.com

1997 Bristol Suites ¹
Nack Ballard Mike Senkiewicz
Malcolm Davis
Garry Kasparov IBM Deep Blue
The New York Times

1997

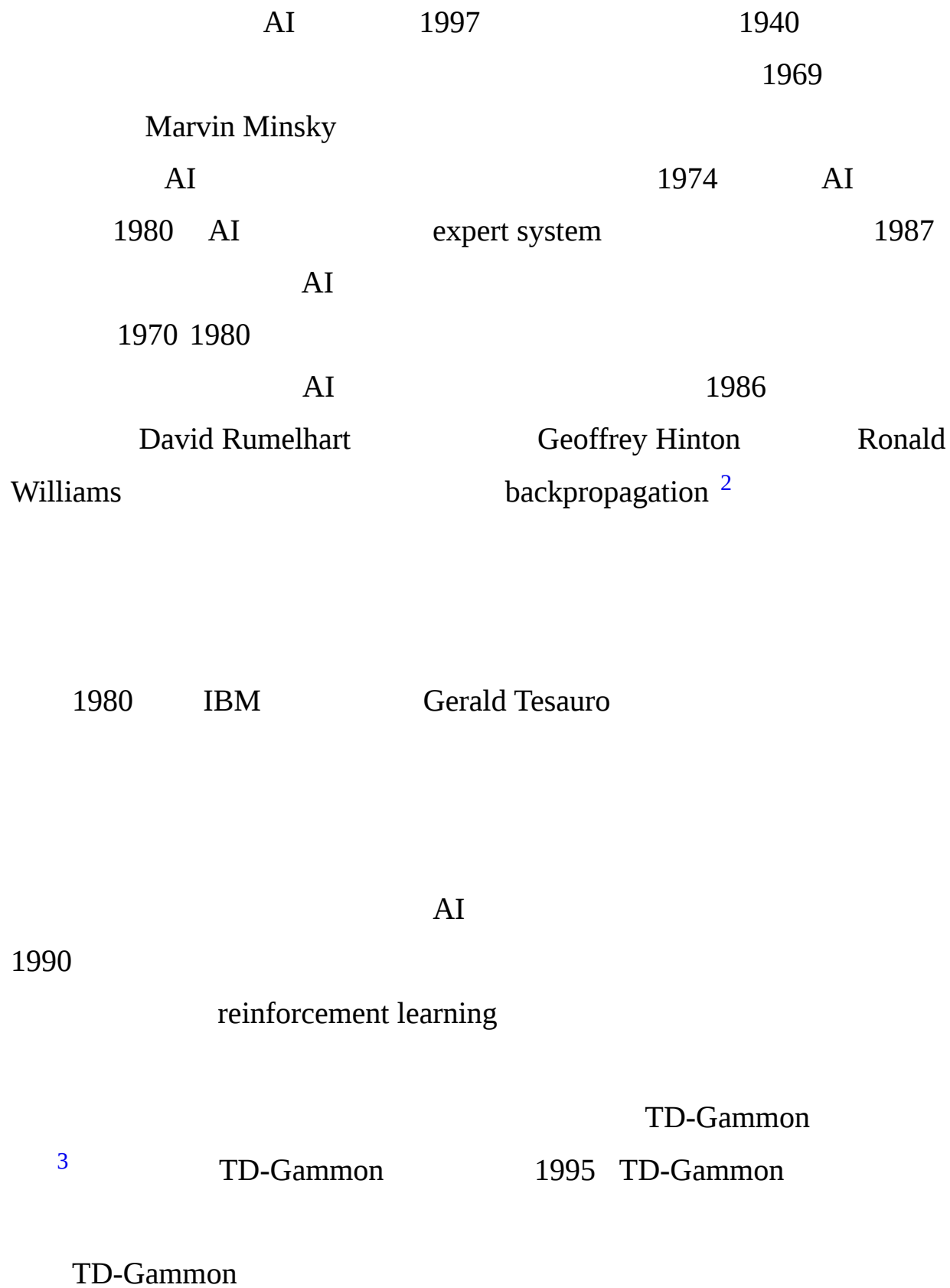
200

AI

Jellyfish

side bet





TD-Gammon

1995

Kit Woolsey TD-Gammon

TD-Gammon

TD-Gammon

• • • • •

IBM

Fredrik Dahl 1994

1983

Matthew Broderick

WarGames

1994

3.5

AI

inference

▲▼▲▼▲▼▲▼▲▼▲▼▲▼▲▼▲▼▲▼▼

11,600

1997

Windows

New Ideas in Backgammon

heads up

limit Texas Hold'em

no-limit hold'em

1970
Werbos

Seppo Linnainmaa

1974
Perceptrons

Paul

1986

3 TD-Gammon Gammon backgammon TD
TD-Gammon

Temporal Difference Learning

OceanofPDF.com

Johnathan Wendel

ID fatal1ty 1999

III *Quake III: Arena*

1980

1996

Wichita

Tiger

Woods

NFL

140 0.14

TNT2

III

140

30 0.03

fatal1ty

2024

TNT2

▲▼▲▼▲▼▲▼▲▼▲▼▲▼▲▼▲▼▲▼▲▼▲▼

1999

6

600

NASDAQ

NVDA 12

60

TNT2

Compaq

S3

Savage 4

Camino

ATI

S3

TNT2

IPO

ATI S3

3dfx

3dfx Voodoo

1999

GeForce

Geometry Force

NV10

3D

2

2,000

GeForce

graphics-processing unit

GPU

GPU

3dfx

benchmarking

GeForce

3dfx Voodoo4

3dfx

3dfx

3dfx

1996 1999

S3

3dfx

3dfx

2000 8 3dfx

Alex Leupp

3dfx

1

3dfx

3dfx

3dfx

3dfx

3dfx

3dfx

3dfx

Darth Vader ¹



3dfx

Ben Garlick

3dfx 3dfx 7,000
3dfx 100

3dfx

100

20

code base

3dfx



11 4.45

1999

Office

Space Mike Judge

3dfx

Karen Huaulme

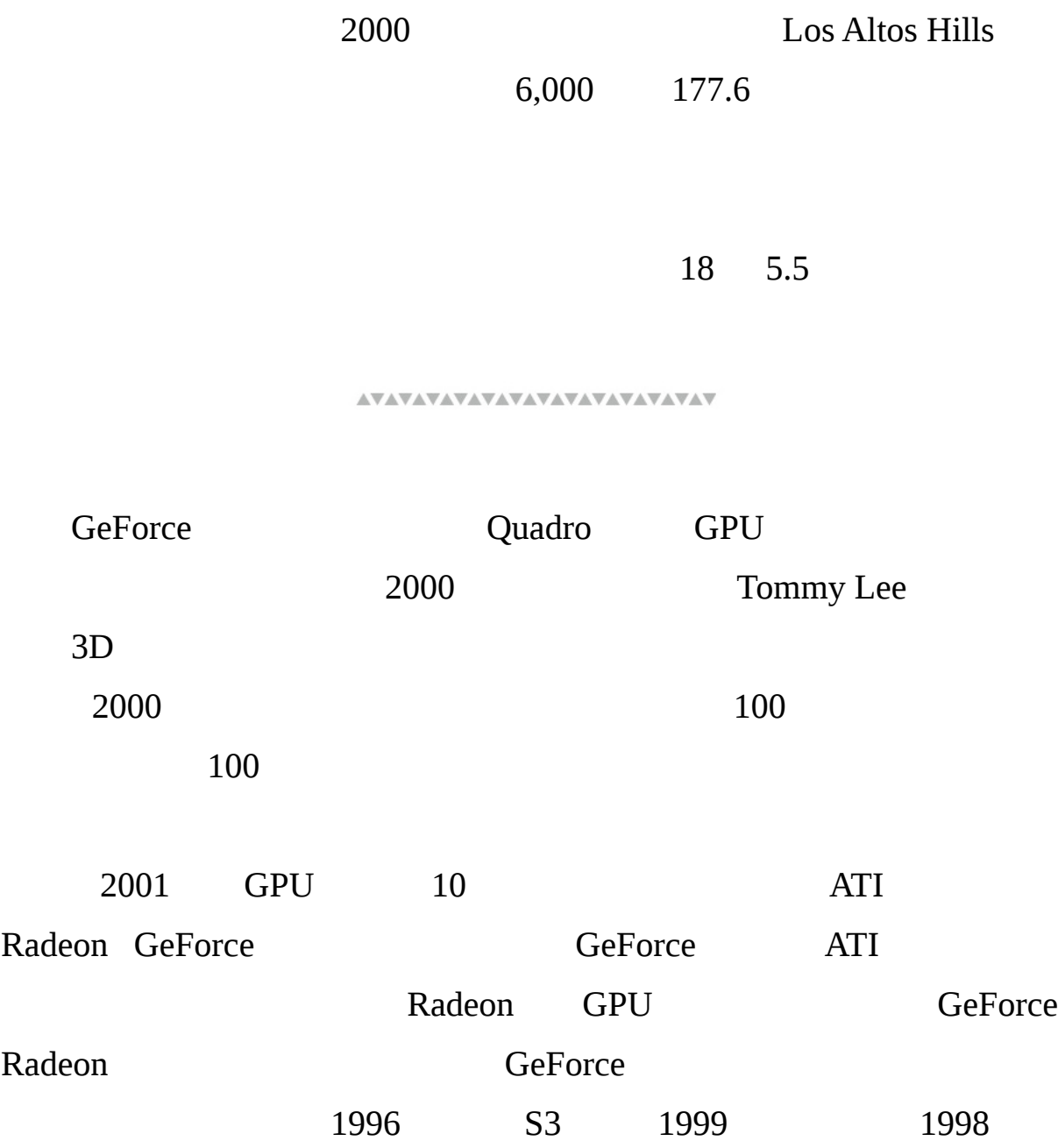
3dfx

iD

T

iD





3dfx

Andy

Grove 10 *Only the Paranoid Survive*

2000

GPU

2001 6 GeForce Xbox 11

Halo

200 500 Enron

500

500

GeForce

Xbox

2002

Harlem Globetrotters

Boeing



2001 2002 90 2002 1
Securities and Exchange Commission Worldcom

Christine Hoberg 1 7

911 500

Bill Gates Xbox
Xbox

Xbox ATI Dawn Cg *The*
Cg Tutorial vertex transformation



2003 GeForce FX

Sharon Clay

Improvement GeForce FX RI Requires
 RI GeForce FX

GeForce

Tim Little



2004 2006

1,000

99.9

It's a Wonderful Life

George Bailey

1,000

5

1,000



Call of Duty

2

Half-Life 2

The Sims

World of Warcraft

2000 2005

4chan

Gamergate

2001

John Hopson

compulsion loop

2004

2000

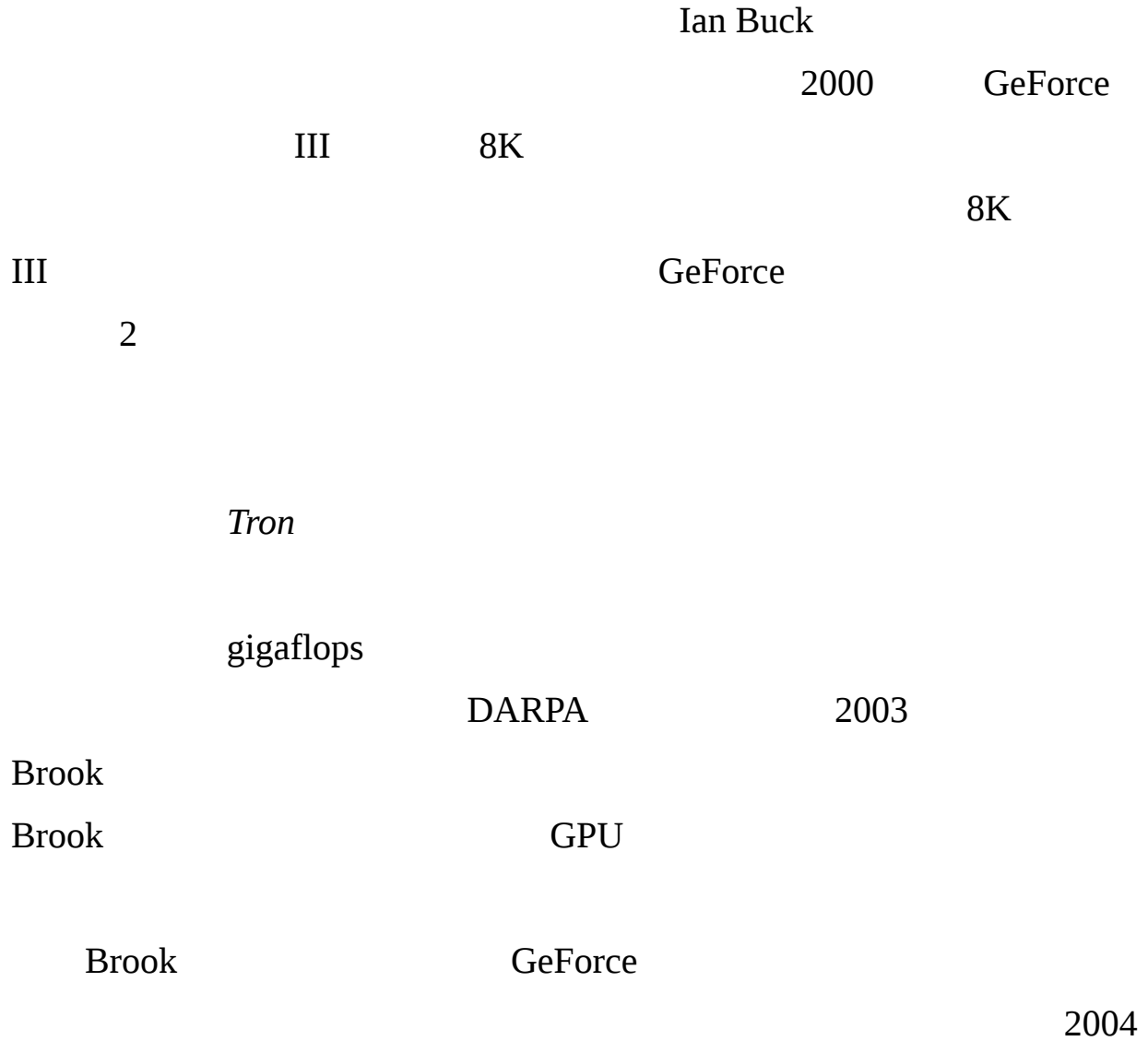
1

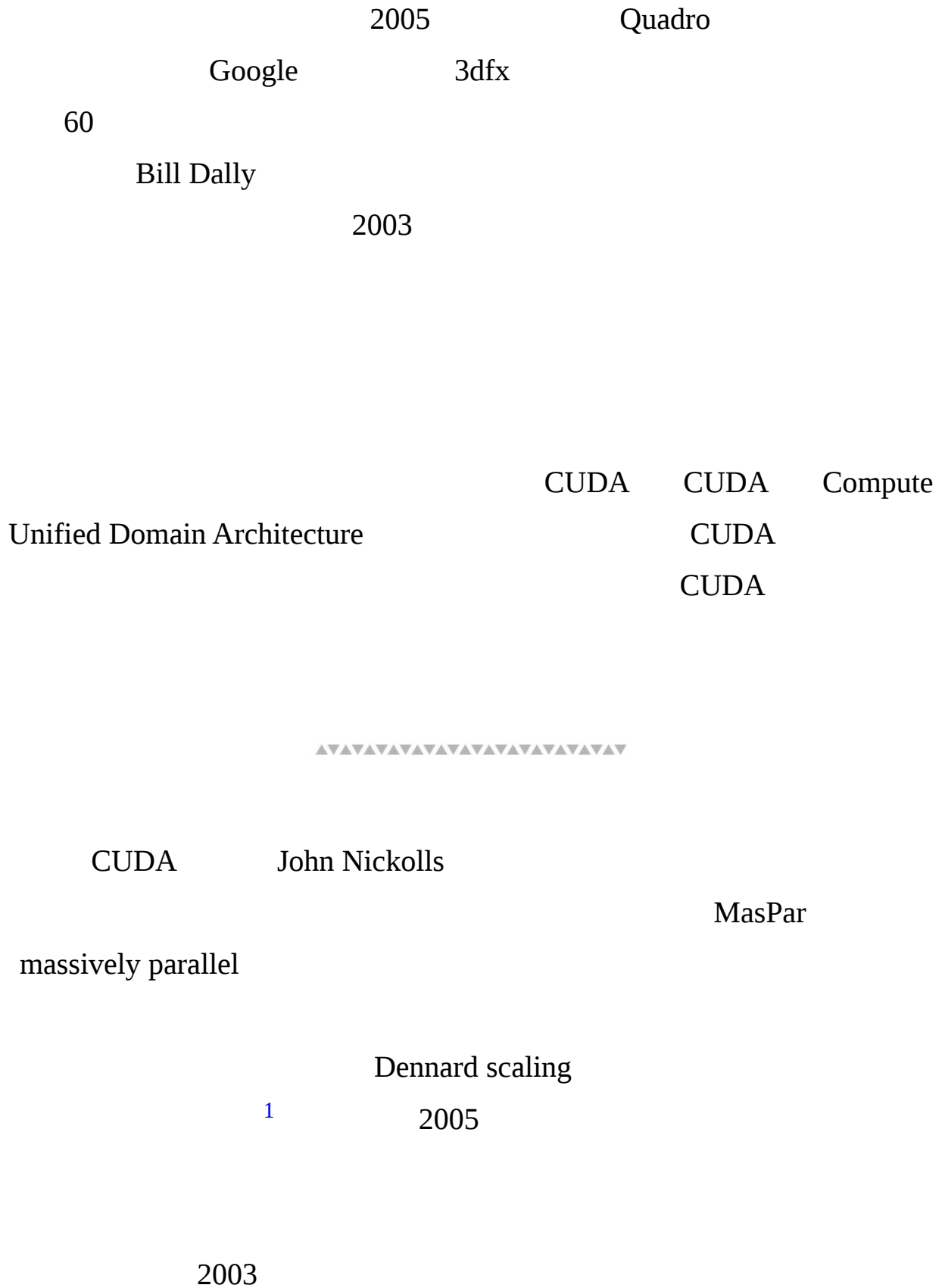
300

60

OceanofPDF.com

CUDA





GeForce

CUDA

Alec

CUDA

CUDA

2000

CUDA

2000

CPU

GPU

Massachusetts

General Hospital

CUDA

zero-billion-dollar

die shot

3dfx

ATI

300



DJ

DJ

DJ

CPU

GPU

CUDA

CPU

GPU

Cuisinart

GPU

GPU

CPU

CPU

CPU

GPU

2000

2000

GPU



CUDA

Arjun Prabhu

GeForce

AMD

Radeon

2

CUDA

CUDA

CUDA

Brett Coon

CUDA

Moscow State University

C++ Python

CUDA

Bas Aarts

CUDA

CUDA

vendor lock

CUDA

CUDA 2006

2007

GeForce

CUDA

CUDA

GPU

thread

CUDA

CUDA

Programming Massively

Parallel Processors

John von

Neumann 1945

General Mills

CUDA

Unreal

Fast

Fourier Transforms

GeForce

Windows

CUDA

2006

GPU

Radeon

2009

2

T

2008

Sameer Halepete



Deb Shoquist

2007

lead time

cycle time

Unit Economics

8 1,000

speed of light

2007

1990

Telluride

Nvidia Research

CUDA

CUDA

10

Lake Tahoe

Grenada

Antigua

Virginia Tech

CalTech

1992

9

Cessna

Long

Island Sound

1,800

Groton-New London airport

Jellybean machine

benchmarking test

2003

2009

2009 2010

Fast Company

2010

Businessweek

2009 1

Jim Plummer



1

GPU

GPU

GPU Technology

Conference GTC 2009

Fairmont Hotel

Kohl's

Woodstock

Mythbusters

Adam Savage

Jamie Hyneman

1,100

GTC

augmented reality AR

Unlocking Biologically-Inspired Computer Vision: a High-Throughput

Approach

Nicolas Pinto

CUDA

GPU

Law & Order DVD

Jerry Orbach

90

CPU

GTC

2010 2011 GTC

2009



CUDA

2000

2011

University of Illinois at Urbana-

Champaign

CUDA

CUDA

2009

2012

CUDA

Podcast Acquired

Ben Gilbert

2012

2013

Starboard Value

Jeff

Smith

CUDA

Darden

Restaurants 6

Olive

Garden



2014

Fortune CUDA

2000 northbridge

CUDA 2010

2011 3.67 Icera cellular

modem Qualcomm

CUDA 2012

CUDA

CUDA

CUDA

Fidelity 1

CUDA

NFL Dawn Hudson



90



GeForce

CUDA

CUDA

17

Elon Musk

CUDA

CUDA

OceanofPDF.com

11

AlexNet

Alex Krizhevsky

2000

University of Toronto

1986

AI

Ilya Sutskever

	CUDA	2008	
Abdelrahman Mohamed	George Dahl		
2009			Neural
Information Processing Systems	NIPS		AI
	NIPS		
	AI		
CUDA	AI		

convolutional neural net

	GPU		
2012		CUDA	GPU
	GPU		
	GPU		

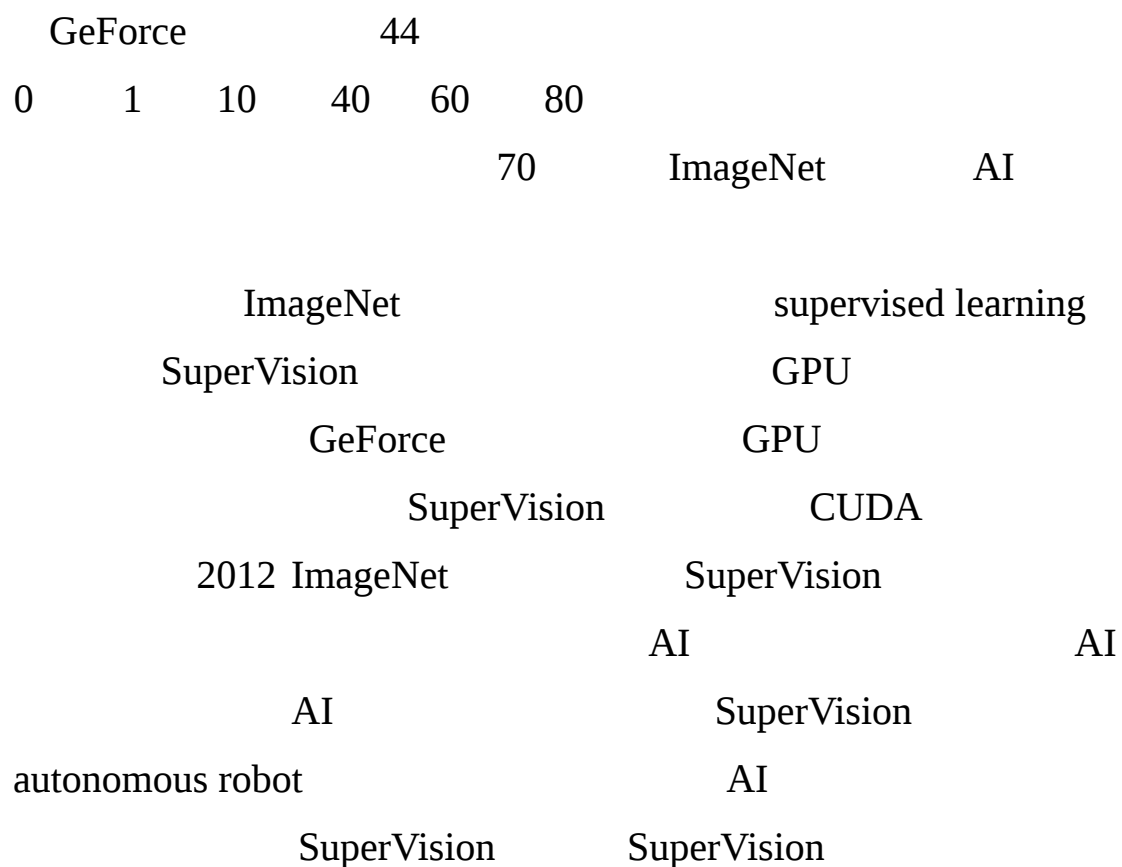
GPU GeForce GTX 580 500 GPU
Alien

GPU

CUDA killer app
 ImageNet
 Amazon Mechanical Turk
ImageNet

CUDA GeForce

synaptic weight



SuperVision ImageNet

2010 2011 2012

80

100

Princeton University

AI ImageNet 2012 10
SuperVision

Uffizi

SuperVision

AI

AI
AI SuperVision
2013 ImageNet 2014

SuperVision

GPU CPU
Google Wojciech Zaremba SuperVision
WojNet SuperVision
Super Vision

AlexNet

AlexNet	DNNResearch	1
DNNResearch		
2012	12	
AlexNet	DeepMind AI	Google
	2,000	
Google	4,400	
Google	AlexNet	

1 DNN Deep Neural Networks

OceanofPDF.com

O.I.A.L.O.

STEM

1

Bryan Catanzaro

T

JPEG

Crime and Punishment

Brigham Young

University

Dostoyevsky

Tolstoy

Pushkin

2001

2000

Parallel

Computing Lab

3D

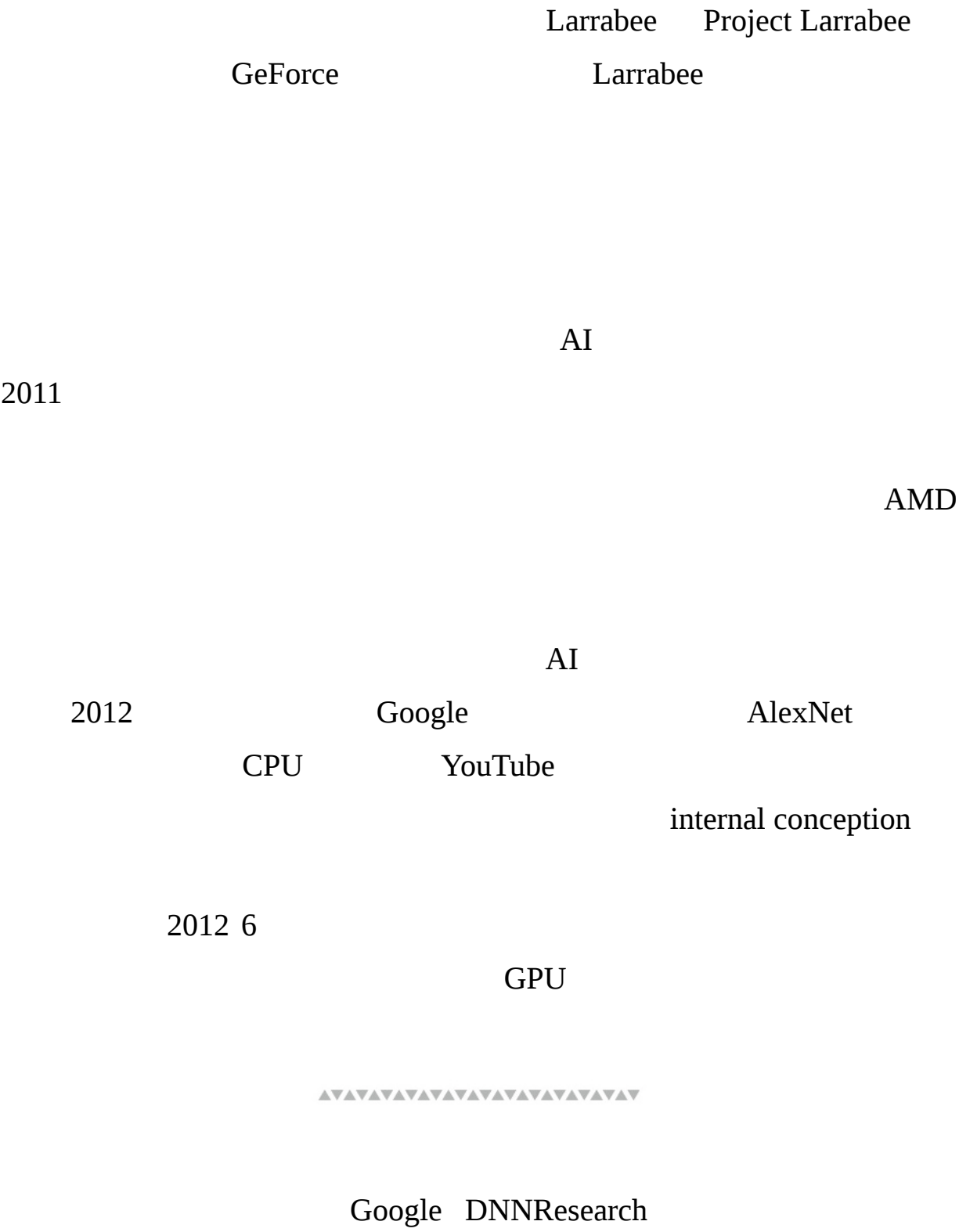
AI

AI

AI

AI

AI



Google

Google

CPU

Google

Google

cuDNN

CUDA

2013

cuDNN

2013 3 GTC

AI

CUDA

2013

CUDA

cuDNN

O.I.A.L.O

Once in a Lifetime Opportunity



CUDA

AI

AI

CUDA AI

Greg Estes

AI



2014

cuDNN

Philippe Vandermersch

CUDA

Fortran

CUDA

AI

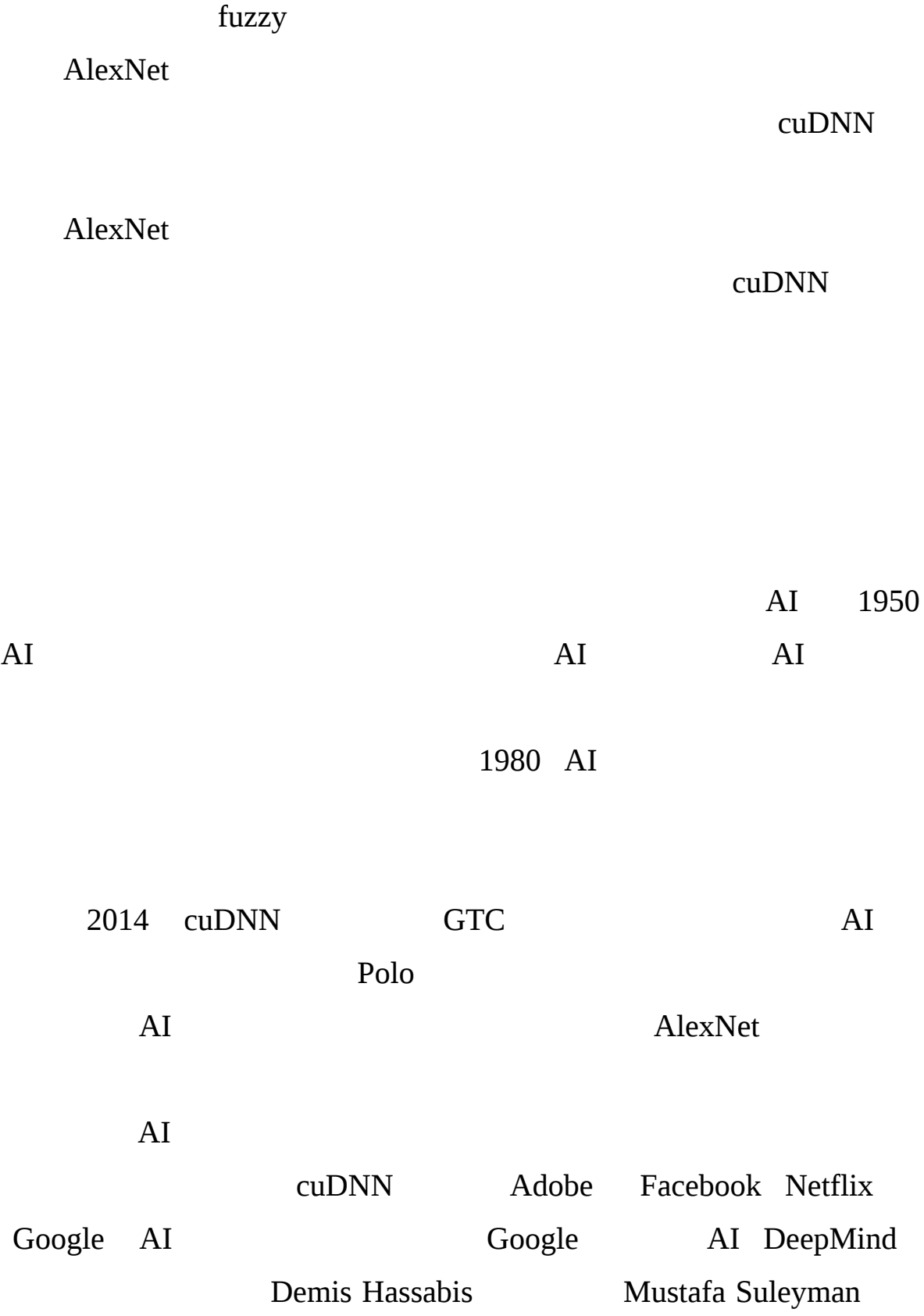
cuDNN

Stephen Wozniak

1

truncation

1



Shane Legg	artificial general intelligence				AGI
DeepMind	AlphaGo	2016			
AI	AI	Google	AlexNet	DeepMind	
Google AI		GPU	2014	Google	
Project Mack Truck			GPU	1.3	
	AI				

DNA

Tesla Roadster Koenigsegg

Maui

Four Seasons Hotel

AI

O.I.A.L.O

Avengers

CGI

Reddit

2014

AI



2013 30 2014 27 2015 66

2001

2001 500

GeForce

AI

Google GPU
GPU

GPU 2 CPU GPU

Curie Tesla Fermi Kepler Maxwell Pascal
Turing Volta Ampere Lovelace Hopper Blackwell
AI

GPU 3 CPU
AlexNet 500
Google Google AI
Google AI
AI Instagram

doomscrolling AI
AI 2010 AI
2015 50 Andreessen
Horowitz Marc Andreessen 2016

2014
CUDA Bas Aarts



AI
Superintelligence

Nick Bostrom 2014

paper-clip maximizer

AI

AI

AI

AI

Edge.org

AI

AI

DeepMind

2015 GTC

Polo

LVMH

2023

Tom Ford

AI

Andrej Karpathy

AlexNet

2011

heads-up display

HUD

AI

AI

AI

AI



2019

social cue

Donald Trump

AI

AI

AlexNet

AI



2015

AI

AI

2015

OpenAI

OpenAI

AI

1.35

4,500

LinkedIn

Reid Hoffman

Y

Combinator

Sam Altman

OpenAI

OpenAI

AI

2015 GTC

Google

AlexNet

OpenAI

North

Dakota

Stripe

Greg Brockman

AlexNet

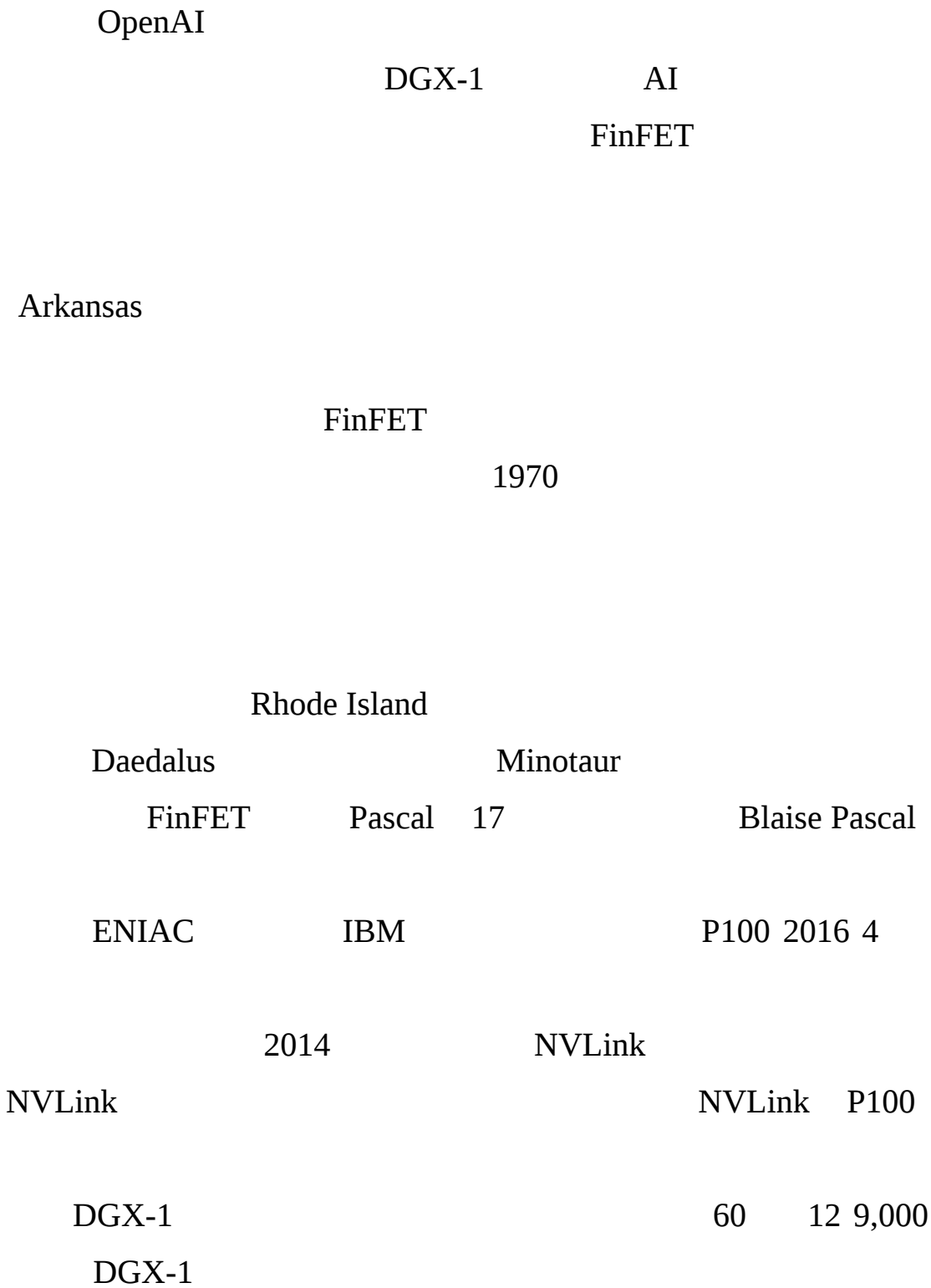
AI

2017

Google

1,500

2019



AI	DGX-1	Colossus	Apple II
	DGX-1	2016 8	DGX-1 OpenAI
Mission district			
		OpenAI	
DGX-1			
	DGX-1	2016	ImageNet
		98	95



AMD AMD AMD 1984
AMD 2006 AMD
AMD ATI AMD
2008 AMD The Summit at Lantana
AMD AMD 2014
2
45 AMD

AMD CPU
AI AMD 2016 Radeon CUDA
CUDA AMD GPU

AMD AMD
ATI GameCube Wii Xbox
Nintendo Switch

Walt
Disney

Switch

Switch

ATI

ATI

Tegra

Switch

Tegra

CPU GPU

Tegra

Switch

Switch

2017

Switch



2017

Observatory

2017

Laser Interferometer Gravitational-Wave

2,900

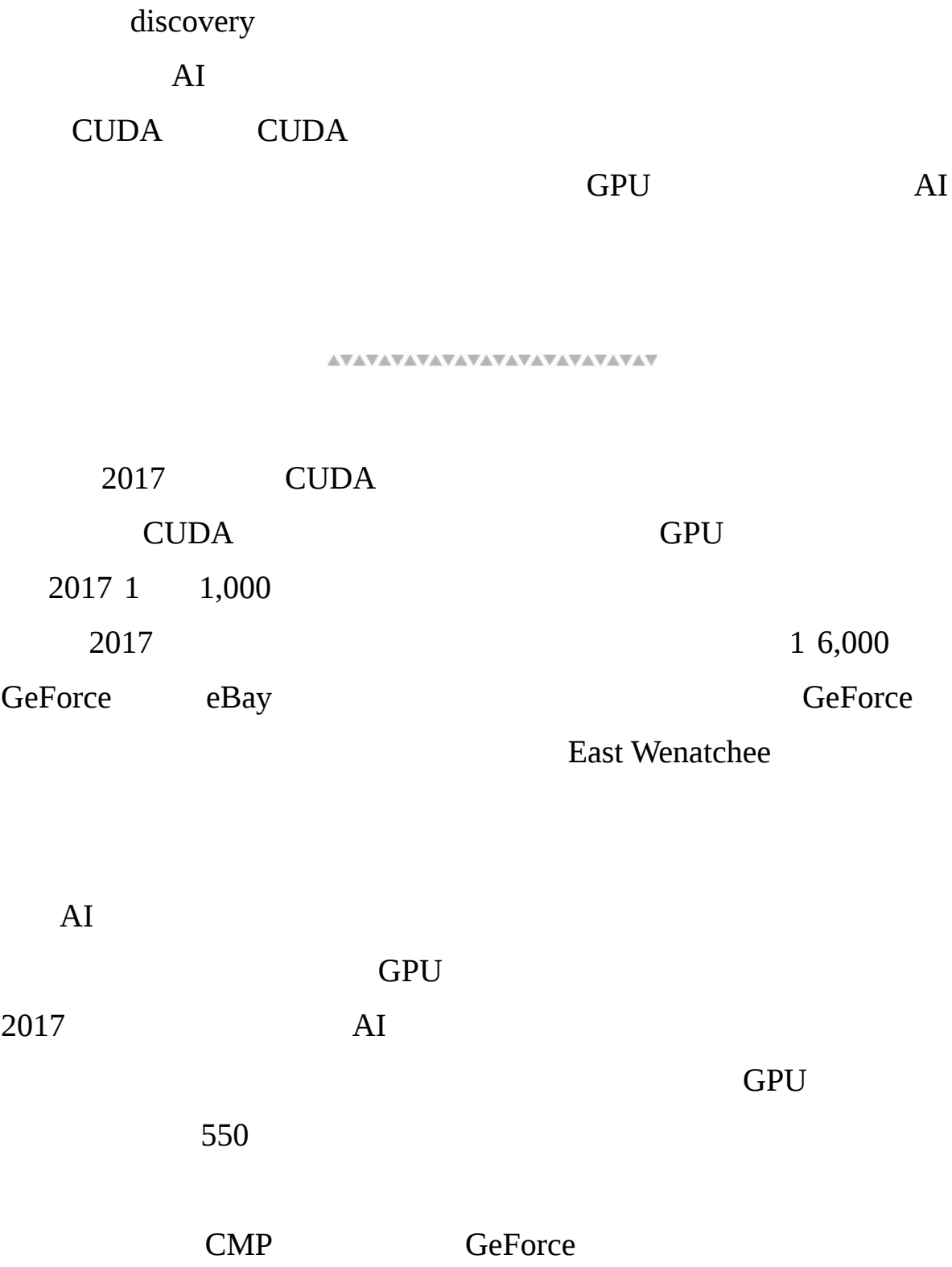
amorphous ice

GPU

GPU

GPU

Science



2018

GPU



2017 1

10

2001

Endeavor

2017 1

Hao Ko
Gensler

Gensler
junior architect

GPU

Loop
David Shannon

Google GooglePlex Infinite
Shannon's

Endeavor Voyager

3D
NV

Voyager



2017 CUDA

CUDA AI

2012

2017 CS 229

CUDA

Now

Azure 2017

Amazon Web Services 50

Satya Nadella

2017 GeForce

GeForce

GeForce Now

Chebotar USC Yevgen
Wayne Gretzky

Isaac Isaac Omniverse
industrial metaverse digital twin
Omniverse Omniverse



2017 30 1,000
2017 2017
AI AI AI

2013
CUDA Switch

1

Star Trek

OceanofPDF.com

15

Transformer

Jakob Uszkoreit

computational linguistics

Hans

2008 Google

Google

recurrent neural network RNN

2014

1 2023

knowledge graph

ness unhappiness token GPU 2023 Google Illia Polosukhin Google Denis Villeneuve 2016 Amy Adams heptapod Arrival Google Ashish Vaswani 2017 Google autobot transformer ¹ 2017 2 transformer Noam Shazeer Llion Jones transformer Google transformer

transformer

transformer

Aidan Gomez

transformer

transformer

transformer

beam search

transformer

transformer

The animal didn't cross the street because it was too tired.

The animal didn't cross the street because it was too

wide.

transformer

tired

wide

an

old Canadian maple tree

a maple Canadian old tree

transformer

AlexNet

transformer

transformer

GPU

ablation

transformer

transformer

transformer

transformer

transformer transformer
transformer
AI 2017 *Neural Information*
Processing Systems AlexNet The
Beatles Attention Is All You Need

2017 7 Lukasz Kaiser transformer
transformer
Transformer Transformer

transformer Google
Google

transformer 2023 transformer Google
Character.Ai Google 2024 Google
Google
Google Google transformer



DGX-1 OpenAI
Dota 2 OpenAI

2018 2 Y Combinator
OpenAI AI AI
The Wall Street Journal
AlexNet
AI
transformer OpenAI
AI transformer
transformer
transformer
pretrain
Transformer Generative Pre-Trained Transformer GPT
GPT-1 2018 6 BookCorpus
Twilight GPT
Dadaist GPT-1 OpenAI
OpenAI DeepMind Google 0
OpenAI Google OpenAI
GPT-1 2012
GPT-2
GPT-2 OpenAI
Andes Mountains
GPT-2

GPT-2			GPT-2
	zero-shot		<i>The Origin of Species</i>
	Charles Darwin	GPT-2	
Tesco	<i>Terminator</i>		John Connor
	Arnold Schwarzenegger		
Tiber			
GPT-2		AI	emergence
GPT-2		1 10	
transformer	AGI	AGI	
GPT	OpenAI	AGI	AI
		AI	AI
AGI			
	AGI		AI
	AlexNet	transformer	
AGI			
GPT-1	GPT-2		
OpenAI		2019	capped-
profit	100	OpenAI Global LLC	
AGI			
	OpenAI 10	1,000	

1,000 GeForce
AGI OpenAI
AI

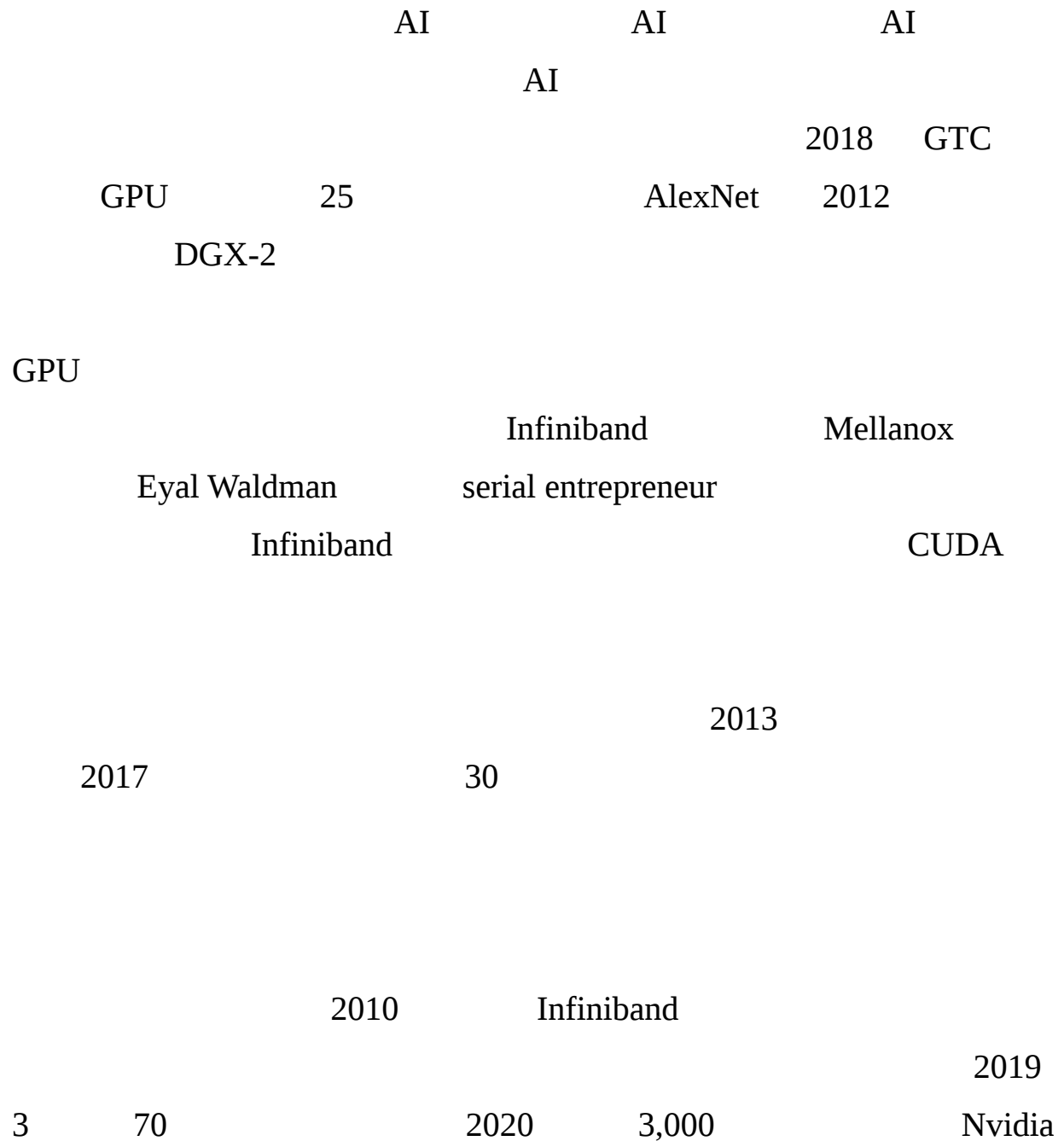
1 autobot

Transformer

transformer

[OceanofPDF.com](https://www.oceanofpdf.com)

16



Israel

AI



2020 9 400 ARM

2021

AMD ARM

AI 90 Google

AI Cerebras

Cerebras Andrew Feldman

AMD CUDA

AI AMD

CUDA

IBM

X

GPU

Eternal Sunshine of the Spotless Mind

2012 2022 AI

AMD

Drive BioNeMo

Clara Morpheus Highlights

Fortnite

2020

treasures

2024

H100

CUDA



GTC

2020 7

AMD

2020

	2:00	2:05	6:00	6:05
	transformer		OpenAI	Azure
GPU	500			
	transformer			RNN
transformer				RNN
	transformer			
	motion-transfer		AI	<i>Mission</i>
<i>Impossible</i>		Tom Cruise	
	2019			
AI	PicassoGAN			PicassoGAN
	AI	2018		

2020 OpenAI GPT-3 1TB
 OpenAI fair use
 1851
 Books2 Books2 LibGen
 GPT OpenAI Jonathan Franzen
 John Grisham Jodi Picoult George R. R. Martin
Times OpenAI
 GPT-3 GPT-3
 2022 ChatGPT OpenAI
 OpenAI
 ChatGPT YouTube

 ChatGPT 2022 11 30
 ChatGPT

 2023 1 ChatGPT 2023 3 OpenAI
 GPT-4 OpenAI GPT-4
 AP 5 GRE 1
 8 GPT-4 GPT-4
 1990 iPhone GPT-4

VGA

GPT-4

ChatGPT

ChatGPT

ChatGPT

ChatGPT

ChatGPT



OpenAI

1

GPT-4

GPT-3

GPT-4

GPT-4

emotional valence OpenAI GPT-4

GPT-4

0.25

GPT-4

10

OpenAI

GPT-4

20

2023 3

transformer

AI

Copilot

DeepMind AlphaFold2

RNA

0 1

AI

OpenAI DALL-E

Midjourney Stable Diffusion

Udio Suno

Jasper

OpenAI 2024 Sora

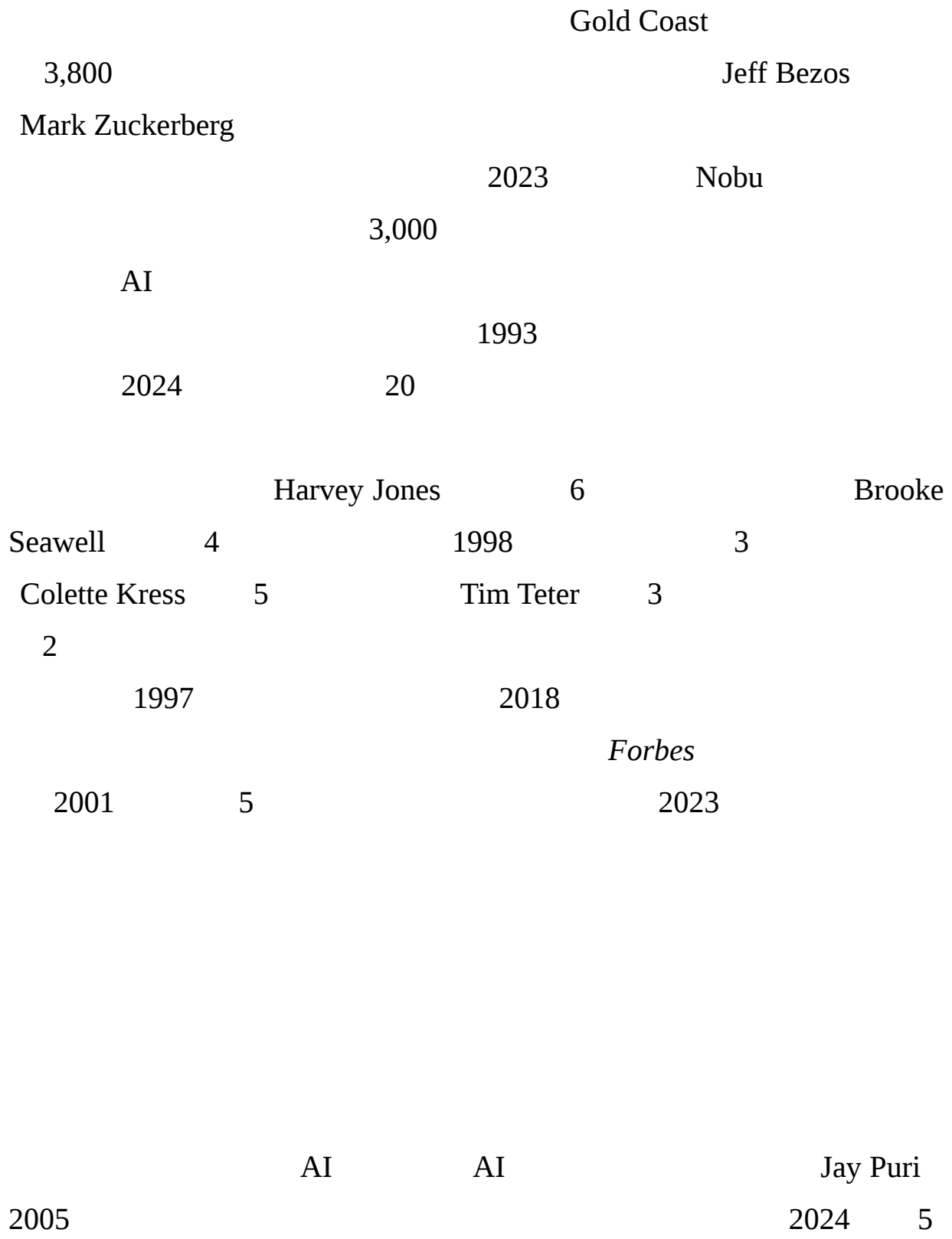
AI

AI

AI AI 2022 DeepMind
AI
AI AI voice-cloning Joe
Biden *Overwatch*
deep-fake
2,500 AI
AI Meta Google
autonomous weapon 2023 AI 500
AI DeepMind
2024 2024 [1](#) National
Academy of Engineer-ing



The Sun Also Rises Mike Campbell
2023 GPU
2023 AI
CNBC 2023 5 25
2,000 Walmart ExxonMobil
AI



OceanofPDF.com

2023

2 1,000

AI

Portal
AI

Jensen

Slack

3

T



2018

ray-tracing

3D

Diane

AI

AI

AI

3D

AI

Book of Genesis

Manhattan Project

AI

AI

AI

AI

AI

George Boole



deep learning super sampling

DLSS

AI

CGI

neural graphics

AI

Google

Burning Man

Arthur C. Clarke

1964 BBC

○○

2001

Prometheus

AI

AI

AI

AI

AI

AI

AI

AI

AI



GPU

Infiniband

[OceanofPDF.com](https://www.oceanofpdf.com)

19

20

100

30

hot stick

Dominion Energy

Loudoun County

Google

AI

2024

AI

AI

OpenAI GPT

Google

AI

TikTok

AI

AI

GPU

GPU

GPU

Google

1 3

watt-hour

AI

Google

GPT

AI

AI

2022

James Earl Jones

2024

MRI

AI

AI

AI

20

GPU

2021

1960

IBM

IBM

70

1980

NASA

2020

2000

Earth-2

AI

AI

2020

A100

250

H100

350

75

2024

B100

700

B200

1,000

2016

DGX

2024

DGX

superpods

Minneapolis

Google

2030

Google

50

2024

Google

Google

AI

OpenAI

AI

AI

OpenAI

IT

GPT

AI

AI

AI

AI

2024
bubble land

Elliott Management
The Financial Times

AI

AI

AI

AI

AI

AI

Sundar

Pichai

[OceanofPDF.com](https://oceanofpdf.com)

20

2024 2 21

CNBC

FinTwit

600

70

46

300

3D

AI

AI

AI

AI

2 22

2,770

650

AI 3

GDP



2022 Glassdoor

Google

IBM

AI

2000

AI

.....

90

2022

2024

2020

CGI

2020

1

LVMH 2020
2022



2023 10 7
Avinatan Or 2024 11

Aviv AI West Bank Tel

2024 3 GTC Blackwell
Run:ai 7
AI 2022
A100 H100

dual-use
ChatGPT

TikTok
2023 6

A100
A100

2024

2024 4

AI

2023

Phoenix 130

Sonora 8 43

Dune

Pyramid Peak 50

Manitowoc 2,300 400

65 Chandler 320

30

3

Pelosi 2022 Nancy
2023



ASML AMD Broadcom Supermicro
2024 3
1999 1 22 2024

2024
Pat Gelsinger 2023 12 GPU AI
1990 3D i740
2000 Larabee 2022
Intel Arc 1

OceanofPDF.com

Jensen

———— Jensen ————

Jensen

Jen-Hsun

2023

Jensen

2011

2024

Jensen

Jensen

Jensen

Jensen

Jensen

Jensen

Jensen

The Manchurian Candidate

Jensen

Jensen



Jensen

2023

Jensen

Jensen

Gensler

Jensen

Jensen

Jensen

AI

shape-shifting

Jensen

Jensen

AI

AI

Omniverse

Gensler

Jensen

Gensler

Jensen

AI

AI

Jensen

AI

AI

Jensen

Jensen

Jensen



Jensen

AI
GTC

2024

NHL

Refik Anadol AI

Ovid *Metamorphosis*

GTC

1964 IBM S/360

AlexNet AI AlexNet GeForce

IBM

podcast

microservice GR00T 1

GPU



2024

Jensen

T

Thomas Burberry

2020

T

1,000

Jensen

3

AMD

4 Jensen

OpenAI

Project Strawberry

AI

Jensen

Fields

Medal

Jensen

[OceanofPDF.com](https://oceanofpdf.com)

22

AI

2023

AI

Yoshua Bengio

AI

2018

A. M. Turing

Award

Yann LeCun

Sephardic Jew

1970

1990

AlexNet

2010

AI

AI

AI

1

ChatGPT

ChatGPT

2022

AI

AI

AI

ChatGPT

Google

AI

ChatGPT

alignment ¹

AI

AI



1998

AI

AI

AI

AI

AI

AI

17

AI

AI

AI

AI

AI

AI

AI



AI

p

AI

p 0 p 50 0
2023 p 50 AI AI
2024 10 20
p 0
AI
AI 1 SB 1047
Google
AI SB 1047 80 2024 9
Gavin Newsom
SB 1047 AI
AlexNet transformer AI
AI
AI 2023 11 OpenAI
OpenAI
AI 2023 9
AI
AI
2024 OpenAI 2024 5 OpenAI
GPT4o o omni AI
GPT4o GPT-4o

OpenAI
OpenAI GPT-4o
Her AI OpenAI GPT-4o Google augmented-
reality AR AI Astra Anthropic
AI Claude GPT-4o
AI sovereign AI xAI GPU
2015 AI Meta 300

GPU

AI

AI

AI



AI ChatGPT
GPT-4 100 78 77

Luddite iPhone

AI
AI

The Book of

Ecclesiastes

AI

AI

AI

AI

AI

self-sustainability

AI

2024

AI

Deep Utopia

2024 GTC

Behavior-1K

Omniverse AI

AI

AI

AI

AI

AI

2024 GTC

GPU

10

GPU

GTC

Robert

Oppenheimer

60

60 Minutes

Bill

Whitaker

Federal Trade Commission

FTC

Lina Khan

AI

OK

.....

CUDA



Eos Eos
Marc Hamilton

OpenAI GPT-3

AI Omniverse

angstrom

CUDA

Eos

GPT-4

3×10^{25}

[OceanofPDF.com](https://oceanofpdf.com)

Bob Sherbin

Willing Davidson

Viking

Allison Lorentzen

Bodley Head

Stuart Williams

Sean

Lavery

Yinuo Shi

Anna Kordunsky

Chris Parris-Lamb

Meghan McEnery

Jay Budzik

Marcus Moretti

Leonard

Diana

Emily

OceanofPDF.com

AI

AI

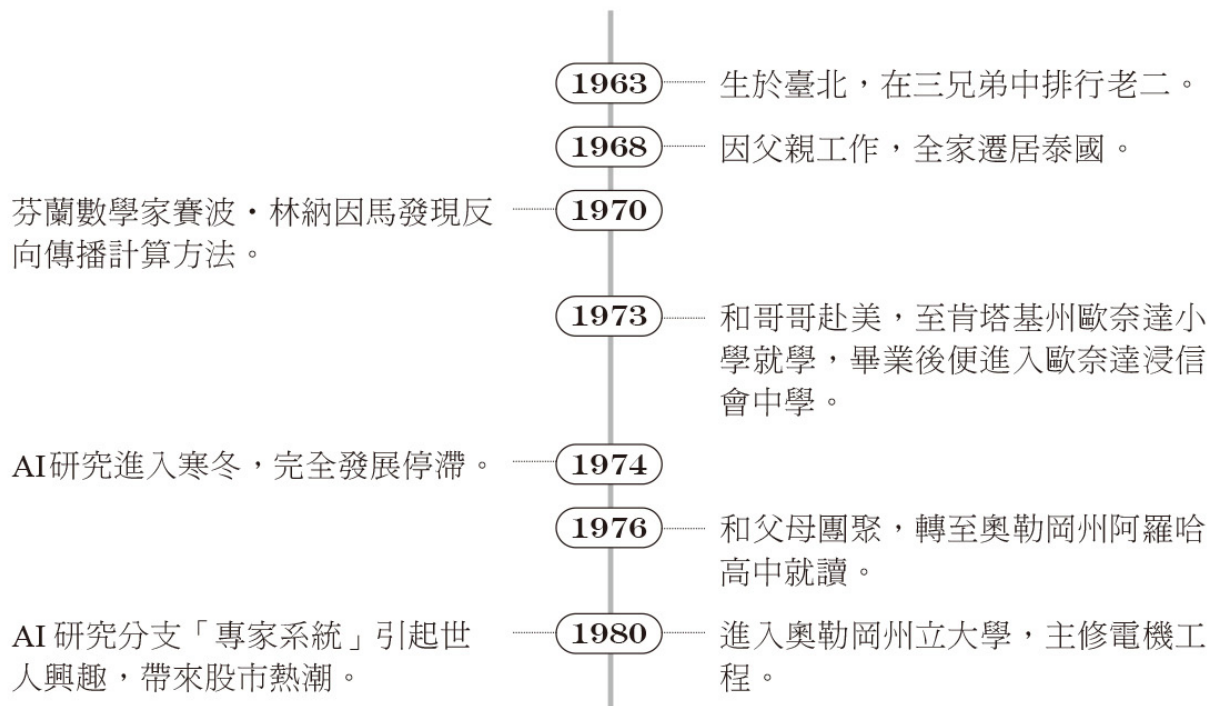
AI

AI 大事記

黃仁勳與輝達大事記

AI 研究的無盡寒冬與黃仁勳的懵懂少年

當黃仁勳對電腦科學著迷時，AI 研究載沉載浮，兩者毫無交集。



AI 研究略有起色，黃仁勳踏入矽谷初試啼聲
AI 領域在各方學術人員努力之下，開始漸有斬獲；
黃仁勳也恰逢其時，在矽革命開始之時踏入矽谷工作

1983 — 進入 AMD，起薪 28,700 美元。

1985 — 轉換跑道至巨積公司，工作之餘攻讀研究所；
結識昇陽電腦晶片設計師克里斯·馬拉考斯基與柯蒂斯·普里姆。

電腦科學家傑佛瑞·辛頓等人提出利用「反向傳播演算法」訓練神經網路，AI 研究重獲新生。

1986

「專家系統」帶來的希望在股災後化為泡影。

1987

和洛麗·米爾斯結婚。

1989

普里姆、馬拉考斯基與黃仁勳合作打造的昇陽 3D 圖形處理器 Sun GX 問世，一炮而紅。

IBM 研究員傑若德·泰薩羅首先成功利用「強化學習」訓練神經網路。

1990

1992

取得史丹佛大學電機工程碩士學位。

AI 技術開始萌芽，黃仁勳展翅創業
當眾人對神經網路棄若敝屣，有人成功利用這項古老的技術賺錢；
黃仁勳帶領輝達度過數次破產危機後，輝達終於被納入標普 500 指數。

1993

和馬拉考斯基、普里姆共同創辦輝達並擔任執行長。

挪威研究員法雷德里克·達爾推出第一個向公眾銷售的神經網路「水母」。

1994

輝達在桑尼維爾租下一間破舊的辦公室。

泰薩羅的神經網路 TD-Gammon 已經不只在學習，甚至開始創新。

1995

輝達第一款晶片 NV1 於秋季上市。

水母和雙陸棋高手打成平手，顯示出新機器時代即將來臨。

- 1996 — 輝達放棄四邊形紋理貼圖技術，NV2 晶片胎死腹中。
- 1997 — 輝達推出 Riva 128，上市四個月銷售百萬張；
開始和台積電合作。
- 1998 — 輝達推出 Riva TNT，以平行運算技術確立品牌定位。
- 1999 — 輝達以每股 12 美元上市，市值約 6 億美元；
輝達推出 GeForce 圖形處理器。
- 2000 — 輝達併購競爭對手 3dfx。
- 2001 — 輝達被納入標普 500 指數，黃仁勳成為全美五百大企業中最年輕的執行長。

AI 領域出現轉機，黃仁勳陷入人生低潮

儘管輝達營收不差，卻接連遭遇不順，得不到市場青睞，
黃仁勳力排眾議重押 CUDA，為正面臨轉機的 AI 研究奠定基礎

史丹佛大學研究生伊恩·巴克與團隊發布開源程式語言 Brook，把輝達的遊戲顯卡變成科學工具。

- 2002 — 美國證券交易委員會調查輝達的財務狀況，當年股價重跌超過 90%。
- 2003 —
- 2004 — 輝達員工數超過千人，營收創歷史新高，股價卻未有起色。
- 2006 — 輝達公開發布 CUDA；
輝達開始傳出產品瑕疵新聞。
- 2008 — 產品瑕疵問題導致嚴重的「焊點門」災情，輝達股價暴跌，市值蒸發近九成。

影像識別資料庫 ImageNet 初版問世，包含超過 1,500 萬張圖片、兩萬兩千個類別。

2009

輝達遭戴爾從遊戲筆電供應商名單剔除；

因「焊點門」問題，輝達上市以來首次虧損；

輝達舉辦第一屆 GTC 大會。

多倫多大學研究生亞歷克斯·克里澤夫斯基，利用輝達顯卡拿下當年的 ImageNet 圖像辨識比賽冠軍。

2012

CUDA 下載量跌到谷底，投資人抱怨連連。

AI 發展步入正軌，輝達乘勢再起

當全世界 AI 人才往矽谷聚集，人們對神經網路另眼相看，黃仁勳沒有錯過機會，輝達只花一個週末便轉型為 AI 公司。

2013

開發 cuDNN，進軍 AI 領域。

Google 收購總部位於倫敦的 AI 公司 DeepMind。

2014

黃仁勳在年度 GTC 大會上宣傳 cuDNN，首次對外提到輝達和 AI 相關；

開始研發資料傳輸高速通道 NVLink。

馬斯克創立 OpenAI，並找來一票傑出的 AI 人才加入。

2015

黃仁勳在年度 GTC 大會上火力全開，展示輝達的 AI 實力，股價漲破 2001 年高峰，一路向上。

ImageNet 圖像辨識比賽標記圖像的正確率達到 98%，突破人類 95% 的平均正確率，機器超越人類的時代就此降臨。

2016

輝達打造出史上功能最強大的電腦 DGX-1，交給馬斯克的 OpenAI 與李飛飛的史丹佛大學實驗室；

輝達打入電玩主機領域：任天堂宣布新一代 Switch 將使用輝達晶片。

Google 研究人員雅寇布・烏斯克瑞特和七位同事打造出自注意力機制軟體 transformer，已經能寫出流暢的文字。

2017

輝達年獲利首次突破 10 億美元，雇用第一萬名員工，市值達 1,000 億美元；
新總部「奮進號」落成；
CUDA 軟體套件的下載次數達到兩百七十萬次；
輝達推出雲端服務平臺。

OpenAI 推出 GPT-1，證明這類大型語言模型確實可行。

2018

OpenAI 推出 GPT-2，生成的文本已經能和人類文本不分軒輊。

2019

輝達收購 Infiniband，這間公司設計出世界上最快的網路連線標準。

AI 領域迎來寒武紀大爆發，輝達坐穩 AI 霸主
transformer 架構和超大規模平行運算相輔相成，造就輝達 AI 霸主地位
華爾街分析師表示：「AI 領域的戰爭開打，而輝達是唯一的軍火商。」

OpenAI 推出 GPT-3，它能解開邏輯問題，也能編寫程式碼。

2020

輝達市值超越英特爾、AMD，成為全球市值最高的半導體公司。

OpenAI 開放聊天機器人 ChatGPT 測試版。

2022

ChatGPT 的每月活躍用戶人數達到一億人；

OpenAI 推出 GPT-4。

2023

輝達 AI 晶片市占率高達九成；
輝達市值大漲約 2,000 億美元，成為全球市值第六大公司，超越沃爾瑪與埃克森美孚的總和。

2024

輝達打敗微軟、蘋果，成為全球市值最高的企業；
自 1999 年初首次公開募股至 2024 年年中，輝達股價上漲超過三十萬倍。

The Thinking Machine: Jensen Huang, Nvidia, and the World's Most Coveted Microchip

— Stephen Witt

—

—

—

—

—

—

—

—

—

—

—

—

—

—

—
— 104 93 1
— 02-2662-0012 —02-2662-0007 02-2662-0009
—cwpc@cwgv.com.tw
—1326703-6

—
—
—
—
— 2517
— —02-8990-2588
—2025 1 20

—
— 130 7 11
— 02 2226-9120
—ebook@iga.com.tw
—www.iga.com.tw

Copyright © [2025 by Stephen Richard Witt]

Complex Chinese edition copyright © 2025 by Commonwealth Publishing Co., Ltd.,
a division of Global Views - Commonwealth Publishing Group

This edition is published by arrangement with The Gernert Company, Inc.

through Bardonia-Chinese Media Agency.

All rights reserved.

—500

ISBN—978-626-417-137-3 EISBN—9786264171359 EPUB

9786264171366 PDF

—BCB868

—bookzone.cwgv.com.tw

OceanofPDF.com