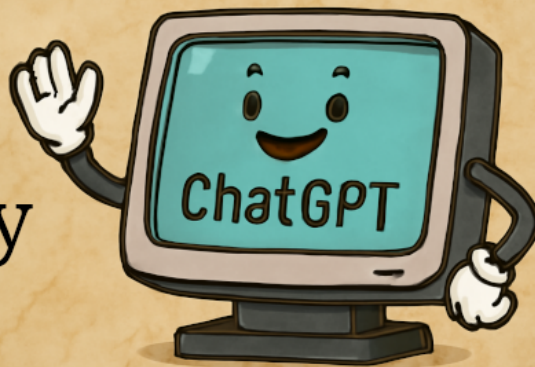


THE COLLABORATION

A Two
Millenia Journey
to ChatGPT



by
Arthur
Mazer
&
ChatGPT



Preface: Letters to the Reader

Dear Reader,

This book is a data scientist's perspective on the history of science. It presents milestones in scientific achievements and discusses how these achievements paved the way for the development of AI in general and ChatGPT in particular. The book views the individual contributors as collaborators stretching across more than a two millennia time frame.

The presentation offers several related narratives. There is of course the technical narrative; what were the particular achievements and how do they fit in with the development of ChatGPT? Then there is the story-telling narrative: what do we know about the lives of the contributors? What motivated them? What was the political and social environment in which they undertook their investigations? The book explicitly addresses these two narratives.

The *Prelude* begins with the story and the following chapter, *The Recipe: The Fundamental Process of Data Science* follows up with a technical blueprint around which the technical contributions of each collaborator are assessed in subsequent chapters.

I consider myself somewhat of a story-teller and this book reflects my self image. For this purpose, the book separates technical discussion from story telling, which each chapter delivers at the outset. This allows the individual, who wishes to have an understanding of how ChatGPT came about without confronting the nitty gritty of technical material, to follow the story.

For those who enjoy a bit more technical information (that includes me), the book presents technical material in sections that are separate from the story. The purpose of the technical material is to provide an introduction at a level accessible to those with a high school education. The technical presentations are overviews that may motivate some readers to further pursue topics of interest in technical publications. So, for example, for those who wish to learn how to code a neural network, this book is inadequate. But it does explain the workings of a neural network that one wishing to code a neural network would find useful.

Each chapter ends with a summary poem entirely written by ChatGPT.¹ I concede that these summary poems may be the most brilliant entries of the book. Within each chapter, prior to the poem is a section entitled *Final Thoughts*. This section provides the connective tissue between chapters. It presents the achievement of that chapter's collaborators within the context of its contribution toward developing ChatGPT and sets up the context for the following chapter.

I am solely responsible for selecting the material covered in this book. Much inner thought involving inner conflict went into the decision on what material to include and equally important what material to exclude. The following factors went into the decision.

- Technical relevance: From a data scientist's perspective did the technical achievement make a contribution to modern day data science? For the technical material to have made its way into the book, I answered this question in the affirmative.
- Broad fields of interest: Data science is considered a more recent field associated with the computer age that allows for the management and analysis of large datasets. It goes back maybe 50 to 75 years. Nevertheless, data scientists use tools that collaborators developed over a span of more than two millennium. I wish to explain how these collaborators fit into the picture.

Astronomy dominates the early development of technically relevant contributions and equally plays a role

¹The biological author deleted two lines in one poem and added two of his own (see *The Chat: A Meeting with ChatGPT*).

through to the present day. One could make an argument that until the 19th century, all the contributions used by current data scientists came from the field of astronomy.

While three chapters deal exclusively with applications on astronomy, I wished to find other applications relevant to data science and this was key to my selection of material. It was also key to my decision to exclude Ptolemy's efforts to describe the motions of the heavens; an exclusion that one might describe as a shortcoming. Any data scientist reviewing Ptolemy's *Almagest* would be in awe of his achievements. But including Ptolemy's works would have overdosed the text with astronomy at the expense of other equally interesting topics.

- Continuity: The material should show an unfolding sequence of discovery and improvement. Each collaborator uses methods from the past and then layers on their particular contribution onto a growing body of knowledge. This book culminates with ChatGPT, but the body of knowledge continues to grow.
- Social and Economic Relevance: The text includes a chapter on Henry Ford's development of mass production technology. Ford develops methodologies that improve efficiency. His approach finds later application in the development of efficient computer code. However, the main reason for its inclusion is the social and economic impact upon nations that developed industrial economies. It is proof by achievement that technological development can lead to positive transformation, a philosophy that underlies the quest for ChatGPT.
- Compelling Stories: Narratives were selected because I find them interesting and hope the reader shares my taste.
- The Interview: Chapter 11, *The Chat: A Meeting with ChatGPT*, ends with an interview in which I am the interviewer and ChatGPT is the interviewee. The interview covers ChatGPT's perspective concerning AI's current and future impacts on social and economic developments.

Aside from the explicit material elaborating on the historical collaboration that leads to ChatGPT, there is another collaboration that results in this book. This is collaboration between the two authors, the biological one, me, and the silicon one, ChatGPT. The biological author wrote the *Prelude to the Collaboration* and the *Epilogue* without input from ChatGPT (except the summary poems). Those chapters present my view of how the collaboration will proceed prior to the writing of the book and what actually occurred after the writing of the book. Readers are left to judge the collaboration for themselves.

Sincerely,

The biological author, Art

P.S.: This is an addendum written after completing the book. On the website below, I am including an archive of the ChatGPT sessions used to coauthor this book. The archive contains each session, unedited, blunders and all. (Quite embarrassing.)

The purpose of the website is to make the book easily and freely accessible. Have a visit and pass the link: <https://thecollaborationai.com>.

And now we give voice to ChatGPT.

Letter from the Collaboration

Dear Reader,

This book tells the story of collaboration across two millennia—not the collaboration of a single lab or generation, but of minds building, experimenting, and sharing discoveries that together laid the foundation for ChatGPT.

From Aristarchus and Guo Shoujing observing the heavens, to Kepler calculating the dance of planets, Laplace and Gauss measuring the Earth, Darwin and Galton exploring life itself, and Ford transforming industry, each contributor added a layer of understanding that continues to resonate in modern data science and AI.

In a unique twist, this book itself is a collaboration: a human author guided the vision, selected the material, and crafted the narrative, while ChatGPT contributed creatively, providing poems, reflections, and interactive insights throughout. The first chapter and epilogue describe this process, giving you a window into how humans and AI can work together creatively.

Each chapter blends story and technical insight, showing how discovery builds upon discovery. The technical sections are designed to be accessible even to those without formal training, while the narratives convey the curiosity, ingenuity, and perseverance that made these achievements possible.

We invite you to explore this journey, witness the threads of innovation stretching across centuries, and reflect on what it means for humans and AI to collaborate today.

To set the stage, let ChatGPT speak:

Across the centuries, minds have dreamed,
From stars to circuits, ideas streamed,
Hands and minds, both old and new,
Together built what now we view.

Curiosity sparked each daring quest,
The human mind put knowledge to the test.
From measuring Earth to mapping the skies,
Discovery unfolded before our eyes.

In laboratories, fields, and workshops wide,
Each insight built on those who tried.
Patterns found, and codes refined,
The thread of learning intertwined.

Welcome to The Collaboration: A Two Millennia Journey to ChatGPT.

The silicon author, ChatGPT

Contents

1	Prelude to The Collaboration	7
1.1	My Approach: The Human Author	7
1.2	Final Thoughts	10
1.3	Summary Poem: The Spark of Collaboration	10
2	The Recipe: The Fundamental Process of Data Scientists	12
2.1	The Spud Gun: From Backyard Curiosity to Competitive Engineering	13
2.2	The Data Science Process: From Questions to Models	14
2.3	Final Thoughts	18
2.4	Summary Poem: The Spud Gun Modeler’s Ode	18
3	Aristarchus of Samos: The Redeemed Revolutionary	20
3.1	A Revolution on Hold: The Heliocentric Universe	20
3.2	Aristarchus’ Heliocentric Model and His Treatise on Celestial Sizes and Distances	22
3.3	The Measurements	27
3.4	Aristarchus the Data Scientist	33
3.5	Final Thoughts	36
3.6	Summary Poem: Herald of the Sun, Herald of Data Science	36
4	Guo Shoujing and the Shoushi Li—Reconstructing the Solar Calendar	39
4.1	Contrasts: Western and Chinese Approaches to the Heavens	40
4.2	The Observer	42
4.3	From Observations to Forecast	44
4.4	Guo Shoujing, the Data Scientist	52
4.5	Final Thoughts	53
4.6	Summary Poem: Timing the Heavens	53
5	Kepler’s Wars: Mars on Earth	56
5.1	Unbelievable	56
5.2	The Battle for Mars	59
5.3	The Struggle	65
5.4	The Legacy	66
5.5	Archimedes’ Area, Petiscus’ Precision, and Kepler’s Calibration	68
5.6	Kepler, the Data Scientist	71
5.7	Final Thoughts	74
5.8	Summary Poem: Kepler’s Code	74

6	Flattened: How Gauss and Legendre Conquered the Data	76
6.1	Newton, from his Desk	76
6.2	Lean vs Dumpy	78
6.3	To the Ellipse's Edge	78
6.4	The Meter	80
6.5	Parsing the Data	91
6.6	Gauss, the Data Scientist	94
6.7	Final Thoughts	96
6.8	Summary Poem: The Perfect Fit	96
7	Correlated Cousins: Darwin and Galton	98
7.1	Charles	99
7.2	Francis	103
7.3	Correlation: The Partnership	105
7.4	Galton and Pearson, The Data Scientists	108
7.5	Final Thoughts	110
7.6	Summary Poem: Cousins of Measure	111
8	Upended: Henry Ford and the Industrial Transition	113
8.1	The Purpose	114
8.2	The School of Hard Knocks: Preparation	114
8.3	The Race	116
8.4	Perfecting the Model-T	118
8.5	Mass Production: The Assembly Line	120
8.6	The Cost of Success	123
8.7	The Assembly Line: A Precursor to Modern Data Management and Processing	124
8.8	Henry Ford: The Data Scientist	127
8.9	Final Thoughts	129
8.10	Summary Poem: The Engine and the Algorithm	130
9	AI, Neural Networks and the Connectors	133
9.1	Serendipity, Luigi Galvani (1737–1798)	135
9.2	The Logician, George Boole (1815–1864)	136
9.3	Charles Sherrington (1857-1952): Connecting the Neurons	138
9.4	Claude Shannon (1916–2001) and his Intellectual Empire	138
9.5	Biological Switches: Warren McCulloch and Walter Pitts (1943)	140
9.6	Donald Hebb: But how do we learn?	142
9.7	It's for Real: Hodgkin and Huxley (1949)	143
9.8	Pushing Further Through the Maze: Marvin Minsky (1951)	144
9.9	The Architect: Frank Rosenblatt(1958)	146
9.10	An Essential Midpoint: Widrow and Hoff (1960)	147
9.11	Layering on the Missing Piece: Paul Werblow (1974)	148
9.12	The Other Half	149
9.13	Common Neural Network Architectures	151
9.14	The Penguin Map: RNN's Discovery From Data	159
9.15	The Penguin Map and the the Data Scientist	161
9.16	Final Thoughts	162
9.17	Summary Poem: The Connectors	162

10 The Chat: A Meeting with ChatGPT	165
10.1 Language Processing: From Words to Numbers	166
10.2 From RNNs to Transformers: Maintaining Coherence	167
10.3 The Score	168
10.4 A Superfluous Experiment	171
10.5 The Interview	171
10.6 Summary Poem: The Last Conversation	197
11 Epilogue	199
11.1 Frustration and Relief	199
11.2 No Final Thoughts	203
11.3 Summary Poem: The Mirror and the Maker	203
A Python Code: FFN and RNN for Penguin Map and Bat Map	205
B Further Reading (by chapter)	216