

目 次

はしがき

序 論..... I

第 1 部 AI とは何か？ 何が問題か？—総論—

第 1 章 AI の歴史と動向..... 8

1 AI とは何か 8
(1) 哲学的・認知的観点からの AI 9 / (2) 技術的・工学的観点からの AI 9 / (3) 学術的・研究分野的観点からの AI 9

2 AI の歴史 10
(1) AI 黎明期 (1940年代～1955年) 11 / (2) ダートマス会議 (1956年) 12 / (3) 第 1 次 AI ブーム (1956年～1969年) 12 / (4) AI 冬の時代 (1970年ごろ～1980年ごろ) 14 / (5) エキスパートシステム (1965年ごろ～1990年ごろ) 14 / (6) 第 2 次 AI ブーム (1980年ごろ～1987年ごろ) 17 / (7) AI 冬の時代 (1987年ごろ～2011年ごろ) 17 / (8) 第 3 次 AI ブーム (2000年代中ごろ～) 19 / (9) 第 4 次 AI ブーム (2022年ごろ～) 20 / (10) ノーベル賞受賞 21

3 AI の現状と展望 21
(1) AI の現状 21 / (2) AI の展望 22

4 考えてみよう 22

第 2 章 生成 AI の社会課題..... 24

1 倫理的・法的・社会的課題 25

2 生成 AI の技術的限界 26

3 生成 AI のリスク 26

(1) 情報面でのリスク	27	/(2) 教育・学習面でのリスク	29	/(3) 労働・経済面でのリスク	30	/(4) 安全・犯罪面でのリスク	32	/(5) 倫理・法制度面でのリスク	33
4 生成 AI の課題への技術的アプローチ	34								
(1) 学習データ	35	/(2) 生成 AI モデル開発	36	/(3) 生成 AI モデルの出力	37	/(4) システム設計	39	/(5) セキュリティ	39
5 考えてみよう	39								
コラム 1 トロツコ問題	41								

第 2 部 AI に何ができて、何ができないか? —技術—

第 3 章 深層学習 44

1 ニューラルネットワークの概要	44								
(1) パーセプトロン	44	/(2) 単層パーセプトロン	45	/(3) 順伝播型ニューラルネットワーク	46	/(4) 畳み込みニューラルネットワーク	48	/(5) 再帰型ニューラルネットワーク	50
2 深層学習	52								
(1) 勾配降下法	52	/(2) 誤差逆伝播法 (バックプロパゲーション)	54	/(3) 深層学習の壁と対策	56				
3 考えてみよう	59								

第 4 章 大規模言語モデルとトランスフォーマー 61

1 言語モデル	61								
(1) 言語モデルとは?	61	/(2) 単語埋め込みと言語モデル	62	/(3) 系列変換モデルと RNN	63	/(4) トランスフォーマーによる系列変換モデル	65	/(5) トランスフォーマーによる大規模言語モデル	68
2 大規模言語モデルの概要	69								
(1) 事前学習と言語モデル	69	/(2) ファインチューニングと文脈内学習	69	/(3) 大規模言語モデルのチューニング	71				

3 大規模言語モデルとデータセット	72
4 考えてみよう	73

第5章 説明可能な AI 77

1 説明可能な AI とは	78
(1) 説明可能な AI の概念	78 / (2) 説明可能性と説明責任 78 /
(3) 説明可能性と精度	79
2 説明可能な AI の手法	80
(1) 手法の分類	80
3 代表的なポストホックな説明手法	81
(1) Grad-CAM	82 / (2) LIME 83 / (3) SmoothGrad 84 /
(4) SHAP	85
4 GUI による説明可能な AI ツール	86
5 説明可能な AI による精度向上	87
(1) アテンション機構による精度向上	87 / (2) モデルへのデータ影響度 88
6 不確実性の評価	89
7 生成 AI による説明	90
8 説明可能な AI の課題	92
(1) 説明手法の選択	92 / (2) 計算コスト 92 / (3) ランダム性 92 /
(4) 羅生門効果	93 / (5) 「説明可能な AI」と因果関係 93 /
(6) 法的適用への課題	94
9 考えてみよう	94

第6章 機械学習における公平性 96

1 責任ある AI (Responsible AI)	97
(1) 公平性 (Fairness)	97 / (2) 説明責任 (Accountability) 97 /
(3) 透明性 (Transparency)	98
2 公平性の考え方	99
(1) 公平と平等	99 / (2) 機械学習で扱う公平性 99 / (3) グループ間公平性と個人間公平性 100 /
(4) 間接的公平性	101 / (5) バイアス 102

- 3 機械学習における公平性の扱い 103
 - (1) 混同行列 103 / (2) センシティブ属性の取扱い 104 / (3) 公平性の定義 105 / (4) バイアス緩和手法 109
- 4 考えてみよう 111

第3部 社会問題としてのAI—法律—

第7章 AIと著作権法……………114

- 1 知的財産法 115
- 2 特許法と著作権法 116
 - (1) 保護対象 116 / (2) 独占権の性質 117
- 3 著作権と所有権 118
- 4 著作物 119
 - (1) 著作物の定義 119 / (2) プログラムの著作物 120 / (3) 学習済みパラメータ（重み）の著作物性 120 / (4) AI生成物の著作物性 121
- 5 著作者の権利 122
 - (1) 著作者人格権 123 / (2) 著作権 123
- 6 著作権の制限 124
 - (1) 権利制限規定 124 / (2) AI学習データ 125
- 7 著作権の取引 126
 - (1) 譲渡 126 / (2) 利用許諾 127
- 8 著作権の権利侵害 127
 - (1) 著作権侵害の救済 127 / (2) 依拠性と類似性 128 / (3) 作風 130 / (4) 侵害主体 130
- 9 米国著作権法でのフェア・ユースの抗弁 130
 - (1) フェア・ユースの法理 130 / (2) フェア・ユースに関する重要な判例 131 / (3) フェア・ユースの抗弁 133
- 10 考えてみよう 135

コラム2 パブリシティ権 136

第8章 AIと特許法.....138

- 1 特許権とは 138
- 2 特許権を得るには 139
 - (1) 発明 139 / (2) 特許権を得る条件 141
- 3 特許権の内容 146
- 4 AIと特許 146
 - (1) ソフトウェアと特許 146 / (2) AIと特許 147
- 5 考えてみよう 152

コラム3 AIは発明者になれるのか 154

第9章 データ保護法制とAI開発.....156

- 1 AI学習のためのデータ 157
 - (1) 量と質 157 / (2) データの収集方法 157
- 2 データの法的性質 159
- 3 個人情報保護法 160
 - (1) 制定の背景 160 / (2) デジタル社会での個人情報保護法の改正 162 / (3) 個人情報保護法の目的 162 / (4) 個人情報保護法とプライバシー権 162 / (5) 個人情報 163
- 4 個人情報保護法での事業者が守るべきルール 165
 - (1) 取得・利用段階 165 / (2) 個人情報の利用段階 168
- 5 データ・クローリング 168
 - (1) 「同意」 169 / (2) 本人への通知・公表 169 / (3) 個人情報保護委員会の見解 170
- 6 第三者提供 170
- 7 「EU AI法」での顔認識データベース作成規制 171
- 8 学習済みモデル 173
 - (1) 日本の個人情報保護法 173 / (2) GDPR 173
- 9 EUの個人データ 174
 - (1) EUにおけるデータ保護法制「GDPR」 174 / (2) GDPRにおけるデータ処理の適法化根拠 175 / (3) 特別カテゴリーのデータ 176 / (4) GDPRでの同意リスク 177 / (5) プリュッセル効果 177

- 10 生成 AI への個人情報入力 178
- 11 考えてみよう 179

コラム 4 有害なデータの除去と大規模言語モデルにおける
マシン・アンラーニング 181

コラム 5 マシン・アンラーニングの法的役割 184

コラム 6 忘れられる権利をめぐる日本の判例 187

第10章 オープン・ソース・ソフトウェア (OSS) ライセンス…189

- 1 コンピュータ・プログラムの保護と OSS 190
 - (1) 著作物 190 / (2) OSS と著作権 191 / (3) 「OSS」とプロプライエタリ・ソフトウェア 191
- 2 OSS によるエコシステム 192
- 3 オープン・ソースの定義 193
- 4 OSS ライセンスのランキング 194
- 5 寛容型ライセンスと互恵型ライセンス 194
 - (1) 寛容型ライセンス 195 / (2) 互恵型ライセンス 195
- 6 代表的な OSS ライセンス 196
 - (1) [寛容型] MIT ライセンス 197 / (2) [寛容型] BSD ライセンス 199 / (3) [寛容型] Apache ライセンス 2.0 200 / (4) [互恵型] GPL ライセンス:「コピーレフト」型ライセンス 200
- 7 オープン・ソース・ライセンスでの注意点 202
 - (1) 著作権表示、ライセンス表示 202 / (2) 特許許諾 203 / (3) 自社が秘匿したい技術 203 / (4) ライセンス承継 203 / (5) ライセンス両立性 203
- 8 データサイエンティストの迷い 204
 - (1) ライセンスの選択 204 / (2) ライセンス不明のソースコード 204
- 9 CC ライセンス 205
 - (1) CC ライセンスの種類 205 / (2) 非商用 CC ライセンスと日本での AI 学習 206
- 10 考えてみよう 207

第4部 AIと私たちの社会—AIガバナンス—

第11章 AIの倫理と社会的リスク210

- 1 技術のイノベーションとその社会的インパクト 210
 - (1) AI技術の発展と社会的リスクへの関心の高まり 210 / (2) 技術による社会的インパクトと社会的リスク 212
- 2 社会規範・倫理と技術による社会的リスク 213
 - (1) 人類と社会規範・倫理 213 / (2) 社会的リスク：新技術の利用による社会規範への抵触 214 / (3) 社会的リスクのパターン：社会的、倫理的、法的な課題としての顕在化 215
- 3 AI技術の利用がもたらす社会的リスク 217
 - (1) AI技術の登場と社会的リスクの範囲の拡大 217 / (2) これまでのソフトウェア技術と比較したAI技術の特徴：信頼性確保の必要性 218 / (3) AI技術の利用によって生じる社会的リスクに係るパターン 220
- 4 AI技術による社会的リスクの類型化 222
 - (1) AI技術による社会的リスクの類型化（全体像） 222 / (2) システムの信頼性・生命の安全性に係る社会的リスク（セキュリティを含む） 223 / (3) 人権（プライバシー・公平性など）に係る社会的リスク 224 / (4) コンテンツ関連規範とその社会流通に係る社会的リスク 225
- 5 考えてみよう 227

第12章 AI規制とガバナンス229

- 1 社会的リスクに対応する制度としてのガバナンス 229
 - (1) イノベーションのためのガバナンスの必要性 229 / (2) 社会全体としての技術ガバナンス 230 / (3) 技術と制度（ガバナンス）の共進化と地域的差異 231
- 2 世界におけるAIガバナンスの地域的差異とその類型 234
 - (1) AIに対する将来的不安感と地域的差異 234 / (2) AI規制・ガバナンスに係る類型と分類 236

3	世界と日本における AI ガバナンス政策をめぐる経緯と現状	240
	(1) 第3次ブーム以降の世界の AI ガバナンス政策の流れ (～2022年)	240
	(2) 生成 AI 以降の世界の AI ガバナンスの動向 (2023年～)	242
	(3) 日本の AI 規制・ガバナンス政策の経緯と特徴	244
4	考えてみよう	246

第13章 EU AI 法.....248

1	概要	249
	(1) 立法経緯	249
	(2) 目的	250
	(3) 条文構成	250
	(4) 適用範囲	250
	(5) 適用範囲外	251
	(6) 段階的適用	251
2	リスクベース・アプローチによる規制	252
3	禁止される AI	253
	(1) 規制内容	253
	(2) データサイエンティストが気をつける点	254
4	ハイリスク AI システム	256
	(1) 規制対象	256
	(2) 適用開始	258
	(3) プロバイダの義務	258
	(4) データサイエンティストが気をつける点	259
5	生成 AI への規制	259
	(1) 特定の AI システムのプロバイダとデプロイヤーに対する透明性義務	260
	(2) AI システムの定義	261
	(3) 汎用目的 AI モデルのプロバイダの義務	264
	(4) システミック・リスクを伴う汎用目的 AI モデルのプロバイダの義務	266
6	EU AI 法と欧州整合規格	267
	(1) 欧州整合規格	267
	(2) 行動規範	267
7	考えてみよう	268

コラム7 顔認識 AI と規制 270

第14章 日本での AI 法.....271

1	AI 促進とリスクに対する法律の制定	272
	(1) AI 法	272
	(2) 立法経緯	272
	(3) AI 法の必要性	273
	(4) AI 法の概要	274
2	AI リスクへの国の関与	274
	(1) 指針の整備と問題発生時の調査・指導・助言	274
	(2) 既存の	

	法令との関係	276	／(3) 罰則なきソフトな AI 法	276
3	利活用の促進	277		
4	AI 戦略本部	277		
5	附帯決議	277		
6	施行	280		
	(1) 施行日	280	／(2) 日本流 AI ガバナンス	280
7	考えてみよう	280		

索引