

Contents

Part I Human Nature of Deception and Perceptions of Truth

1	The Problem of Misinformation and Disinformation Online	3
1.1	Introduction	4
1.1.1	Infodemic	4
1.1.2	Mis- and Disinformation	7
1.2	Framework for Causes and Interventions	8
1.3	Immediate Digital Literacy Steps	11
1.4	Terminology in Mis- and Disinformation Research	12
1.4.1	Digital Deception	13
1.4.2	“Fake News”	13
1.4.3	Information Disorder and Other Problematic Content	14
1.5	Computational Fields	16
1.5.1	AI, Computational Linguistics, and Natural Language Processing	17
1.5.2	Shift from Factuality to Subjectivity Analysis	18
1.6	Varieties of Fakes Identifiable with AI	20
1.6.1	Serious Journalistic Fraud, or Falsehoods	20
1.6.2	Rumors or Hoaxes	21
1.6.3	Satirical Fakes	22
1.6.4	Deepfakes	23
1.6.5	Resulting AI Technologies	23
1.7	Conclusions	23
	References	24
2	Psychology of Misinformation and Language of Deceit	29
2.1	Introduction: Deception Is Uncooperative Communication	30
2.2	Motivations to Deceive and Disinform	30
2.2.1	Deceptive Content	32
2.2.2	Serious and Everyday Lies	32
2.2.3	Incentives to Disinform	33

2.3	Many Faces of Deception	36
2.3.1	Definitions of Deception	36
2.3.2	Deception Is Pervasive	37
2.3.3	Deception Varieties and Typologies	37
2.4	Why Mis-/Disinformation Is Effective	40
2.4.1	External Circumstances and Human Propensities	41
2.4.2	Cognitive Biases and Individual Traits	42
2.5	Language of Deceit and Cues to Its Detection	44
2.5.1	Human Inability to Detect Deception	44
2.5.2	Cues to Lie Detection by Mode of Perception	45
2.5.3	Verbal Cues	46
2.5.4	Incentives to Deceive	47
2.6	Putting Psychological Research to Practice: Interventions	47
2.6.1	Fact-Checking and Content Moderation	47
2.6.2	Inoculating	49
2.6.3	Labeling	50
2.6.4	Literacy Education	53
2.7	Summary	54
	References	55
3	Credibility Assessment Models and Trust Indicators in Social Sciences	61
3.1	Introduction	62
3.2	Credibility in Information Retrieval	63
3.3	Credibility Indicators in Social Sciences	64
3.4	Trust Modeling	70
3.5	Credibility Indicators in Journalistic Fact-Checking	73
3.6	Seals of Approval in Libraries and Health Informatics	75
3.7	Composite Reputation Ratings and Popularity Scores	77
3.8	Trust and Distrust Markers in Language	79
3.9	Contemporary Solutions to Credibility Assessments and Trust Ratings of Online Content	80
3.9.1	Digital Realities of Social Technologies in the 2020s	80
3.9.2	“Nutritional” Information Facts for Online Content as Explainable AI	81
3.10	Conclusions	89
	References	91
4	Philosophies of Truth	95
4.1	Introduction	96
4.2	Artificial Intelligence and Applied Ethics (Morality)	98
4.2.1	Moral Virtues	101
4.3	Theories of Truth	104
4.3.1	What Is Truth?	104
4.3.2	The Correspondence Theory	105
4.3.3	The Coherence Theory	108

4.3.4	Pragmatist Theories	109
4.3.5	Epistemic (Anti-Realist) Theories	109
4.3.6	Deflationary (or Deflationist) Theories	110
4.4	Practical Applications of Philosophies of Truth.	111
4.4.1	Ordinary Life Decisions.	111
4.4.2	On Truth, Reality, Fact, and Knowledge	112
4.4.3	Contemporary Debates of “Post-Truth” and “Truth Decay”	113
4.5	Application in AI and NLP	114
4.6	Conclusions	117
	References.	118

Part II Applied Professional Practices and Artificial Intelligence

5	Investigation in Law Enforcement, Journalism, and Sciences	123
5.1	Introduction	124
5.2	Lie Detection in Law Enforcement and Justice System	126
5.2.1	Polygraph (Lie Detector)	126
5.2.2	Verbal Content Analyses	128
5.2.3	Statement Validity Assessment (SVA) and the Undeutsch Hypothesis	129
5.2.4	Content-Based Criteria Analysis (CBCA)	130
5.2.5	Reality Monitoring (RM).	133
5.2.6	Integration of CBCA and RM Approaches	134
5.2.7	Suitability for Automation	134
5.2.8	A Word of Caution.	135
5.3	Investigative Reporting and Fact-Checking in Journalism	136
5.3.1	Journalistic Goals and Aspirations in Pursuit of Facts.	136
5.3.2	Skeptical Way of Knowing in the Best Traditions of Investigative Reporting	139
5.3.3	Fact-Checking within Investigative Journalism and as a Newly Emerged Genre	140
5.3.4	Newsworthiness, Sensationalism, and “Yellow Journalism”	142
5.3.5	News Literacy for all	144
5.3.6	Implications for AI.	145
5.3.7	Lessons Learned from Journalistic Practices.	147
5.4	Scientific Method of Inquiry	147
5.4.1	Scientific Process	148
5.4.2	Scientific Principles	149
5.4.3	Falsifiability	149
5.4.4	Reproducibility (Transparency) and Replicability (Generalizability).	150
5.4.5	Systematicity	151

5.4.6	Lessons Learned from the Scientific Method	151
5.5	Conclusions	152
	References	153
6	Manipulation in Marketing, Advertising, Propaganda, and Public Relations	
6.1	Introduction	157
6.2	Marketing and Advertising	158
6.2.1	Marketing Communications as a Promotional Tool	160
6.2.2	Digital Advertising Formats	161
6.2.3	Ad Revenue Model	164
6.2.4	Advertising Paradigms: Salesmanship in Print, Seduction, and Others	164
6.3	Persuasion in Advertising	166
6.3.1	Classic Modes of Persuasion	166
6.3.2	Advertising Strategies	167
6.3.3	Codes, Standards, and Regulation in Advertising Industry	168
6.3.4	Puffery	169
6.3.5	False Claims and Other Acts of Misleading Advertising	170
6.3.6	Regulatory Actions and Antitrust Investigations	172
6.3.7	Negative Attitude toward Ads, Distrust, and Ad Avoidance	173
6.4	Covert Advertising in Many Guises	174
6.4.1	Irresistibility of Native Advertising	174
6.4.2	Native Ad Varieties: Advertorials, Sponsored Content, Branded Links, and Others	176
6.4.3	Overdue Measures: Regulation, Clearer Labeling, and Awareness of the Masquerade	178
6.5	Virality	180
6.5.1	Elusive Properties of Virality	180
6.5.2	Compelling Word-of-Mouth (WOM)	181
6.5.3	Emotional Vs. Informative Appeals in Advertising	181
6.5.4	Viral Marketing as a Measure of Success in Digital Advertising	182
6.6	Mass Propaganda	182
6.6.1	Consistent and Enduring Efforts to Influence Minds	182
6.6.2	Propagandist Techniques	183
6.6.3	Public Relations	185
6.6.4	Conspiracy Theory Peddling	186
6.7	Implications for Artificial Intelligence and Counterintelligence	187
6.7.1	Virality, Popularity, and Rumor Propagation in AI-Research	187
6.7.2	Ad-Blockers Wars	188

6.7.3	Google's AdWords & AdSense	189
6.7.4	Facebook's Multiplatform Targeted Advertising	190
6.7.5	Bot Wars: Automation of Advertising, Customer Services, and Social Engagement	191
6.8	Broader Regulations and Legislation as Countermeasures	192
6.8.1	Avoiding Governmental Overreach in Freedom of Speech and Citizens' Liberties.	192
6.8.2	Regulatory and Legislative Measures Targeting Social Media Platforms.	195
6.8.3	Regulatory and Legislative Measures Targeting Offenders	195
6.8.4	Regulatory and Legislative Measures Targeting Government Capacity	196
6.8.5	Regulatory and Legislative Measures Targeting Citizens, Civil Society, and Media Organizations	197
6.8.6	Cutting the Finances of "the Disinformation for Hire" Industry	198
6.9	Conclusions	200
	References.	200
7	Artificially Intelligent Solutions: Detection, Debunking, and Fact-Checking	207
7.1	Introduction	208
7.1.1	AI Basic Terms and Concepts Demystified	209
7.1.2	AI Goals in Solving the Mis- and Disinformation Problem	216
7.1.3	Chapter Roadmap: AI Detection by Type of Fake	218
7.2	Automated Deception Detectors	218
7.2.1	Linguistic Feature-Based Approaches to Identifying Falsehoods	218
7.3	Automated Clickbait Detectors	228
7.3.1	Clickbait Defined	228
7.3.2	Clickbait Predictive Features and Clickbait Detection Performance Measures.	228
7.3.3	Sample Clickbait Detector User Interface	232
7.4	Automated Satire Detectors.	235
7.4.1	Satirical Fakes Defined	235
7.4.2	Methodologies for Computing Satire.	236
7.4.3	Predictive Feature Specifications.	237
7.4.4	Methodology for Binary Classification (Satirical Fakes Vs. Legitimate News)	238
7.4.5	Neural Network Methodologies for Computing Satire	240
7.4.6	Sample Satire Detection UIs and Suggested Improvements.	242

7.5	Automated Rumor Debunkers	246
7.5.1	Rumor Detection Stages or Sub-Tasks	246
7.5.2	Rumor Features	248
7.5.3	Other Methodologies for Rumor Debunking	248
7.6	Automated Fact-Checkers	252
7.6.1	Commercial Web-Based and Mobile Applications	256
7.7	Conclusions	257
	References	258
8	Conclusions: Lessons for Infodemic Control and Future of Digital Verification	265
8.1	The Main Premise of the Book	266
8.1.1	Socio-Technological Solutions to Mis- and Disinformation	266
8.1.2	Infodemiological Framework for Causes and Interventions	268
8.2	Concrete Recommendations for Educational Interventions	269
8.2.1	Teaching to Verify your Sources	269
8.2.2	Authoritative Fact-Checking and Debunking	270
8.2.3	Relying on the Scientific Consensus	270
8.2.4	Relying on Professional Verification Experts and Distrusting Manipulators	270
8.2.5	Inoculating and Preemptive Messaging	271
8.2.6	Identifying and Filling Niches in Public Awareness	271
8.2.7	Self-Directed Coaching and Literacy Resources	272
8.2.8	Making Sense of Credibility Indicators, Warning and Labeling Systems	272
8.2.9	The Broader Argument for Informed Rationality and Systematic Training	273
8.3	Automated Interventions	273
8.3.1	AI-Based Tools for Combatting Mis- and Disinformation	273
8.3.2	Claim 1: AI Is Inherently Tied to Human Practices	274
8.3.3	Claim 2: Innovative AI Solutions Are Feasible and in Existence but Are Imperfect	274
8.3.4	Claim 3: Explainable, Accountable, and Ethical AI as the Future of Content Verification	276
8.4	Regulatory and Legislative Interventions	276
8.5	Policy-Making Recommendations for Governments and Institutions	277
8.6	Summary	278
	References	279
	Index	281