
CONTENTS

<i>Preface</i>	v
1. Introduction	1
A (Long) Note on Terminology	6
AI.....	6
AI Decisions.....	8
Developer, User and Producer.....	8
2. Do We Need Fair AI?	10
3. Society's Conception of Fairness	14
3.1. Outcome Fairness.....	14
3.1.1. Merit, Disadvantage, Power and Status.....	15
3.1.2. Collective Outcome Fairness.....	18
3.1.3. Outcome Fairness and AI Development	20
3.2. Procedural Fairness.....	23
3.2.1. Interaction in the Decision Process.....	24
3.2.2. Procedural Fairness and AI Development.....	26
3.3. A Social Fairness Model?.....	27
4. Theoretical Insights from Philosophy and Economics	29
4.1. Philosophy of Fairness	30
4.1.1. AI Ethics.....	30
4.1.2. Fairness and Equality	34
4.1.3. Luck Egalitarianism.....	36
4.1.4. Welfare Egalitarianism	41
4.1.5. Philosophical Guidance for AI Developers.....	43
4.1.6. Philosophy and the Social Conception of Fairness	46
4.2. Economics of Fairness	48
4.2.1. Welfare Economics	48
4.2.2. Measurement, Bias and Incentives	52
4.2.3. Behavioural Economics.....	55
4.2.4. Economics and AI Regulation	57
4.3. Rules and Choices	59

5. Law's Conception of Fairness	61
5.1. Procedural Fairness.....	61
5.1.1. Avoidance of Bias.....	62
5.1.2. Fair 'Hearing'.....	64
5.2. Outcome Fairness.....	67
5.3. Express Outcome Fairness Obligations.....	69
5.3.1. Avoiding Unlawful Discrimination.....	69
5.3.2. Data Protection and Outcome Fairness.....	71
5.3.3. Outcome Fairness in Financial Services.....	82
5.3.4. Framing Outcome Fairness Obligations.....	85
5.4. Legal Fairness Concepts and AI Development.....	87
5.4.1. Common Themes of Legal Fairness.....	88
5.4.2. Lessons for Law Reform.....	92
6. The Human Conception of Fairness	96
6.1. Constituents of Fairness.....	96
6.2. Justifying Unequal Treatment.....	97
6.3. Assessing Fairness.....	100
6.3.1. Fairness Control in AI Development and Use.....	103
6.3.2. AI Fairness Regulation to Meet the Social Conception.....	107
7. AI Fairness Techniques	111
7.1. Machine Learning and Decision Systems.....	111
7.2. Merit Criteria.....	114
7.3. Fairness Assumptions.....	118
7.4. Fairness as Absence of Bias.....	120
7.4.1. Internal Consistency.....	122
7.4.2. Correlation with Human Decisions.....	123
7.4.3. Causality.....	124
7.4.4. Bias and Equality of Opportunity.....	125
7.5. Group and Individual Fairness.....	128
7.6. The Problem of Context.....	129
7.7. Meeting the Human Conception of Fairness Through Regulating AI Development.....	132
8. A Regulatory Framework for AI Fairness	140
8.1. Regulatory Approaches.....	141
8.1.1. Regulatory Principles.....	141
8.1.2. Regulatory Structures.....	144
8.1.3. Regulatory Obligations.....	156
8.2. Aims of the Regulatory Model for Fairness.....	159
8.3. Who to Regulate?.....	160

8.4.	The Obligation to Explain and Mitigate Risks of Unfairness	164
8.4.1.	The Basic Obligation	164
8.4.2.	Sectoral Specifics.....	167
8.5.	Review and Appeal Processes.....	169
8.6.	Scope	170
8.7.	Consequences of Non-compliance	173
8.8.	A General AI Regulator?	175
8.9.	Application in Practice	177
8.9.1.	Bespoke AI.....	177
8.9.2.	Customisable AIaaS System	179
8.9.3.	Generative AI	180
8.10.	Fairness and Wider AI Regulation.....	181
9.	Beyond Fairness	184
9.1.	Ethical AI.....	185
9.2.	Consistent Functionality	186
9.3.	Trustworthiness.....	192
9.3.1.	Technical Controls.....	195
9.3.2.	Human Controls and Accountability.....	197
9.3.3.	Generative AI and Untruths.....	198
9.4.	Explainability	200
9.5.	Use of Intellectual Creations for AI Development.....	207
9.5.1.	Copyright.....	208
9.5.2.	Unfair Competition.....	212
9.5.3.	Regulatory Solutions	213
9.6.	The Future Shape of AI Regulation.....	216
	<i>Select Bibliography</i>	221
	<i>Index</i>	227