

THE GLOBAL LAW COLLECTION

全球法集锦

مجموعة القانون العا

DATA PROTECTION FOR THE PREVENTION OF ALGORITHMIC DISCRIMINATION

PROTECTING FROM DISCRIMINATION AND OTHER HARMS
CAUSED BY ALGORITHMS THROUGH PRIVACY IN THE EU:
POSSIBILITIES, SHORTCOMINGS AND PROPOSALS

ALBA SORIANO ARNAZ

PREFACE

ANDRÉS BOIX PALOP

INCLUYE LIBRO ELECTRÓNICO
THOMSON REUTERS PROVIEW™

THOMSON REUTERS

ARANZADI

DATA PROTECTION FOR THE PREVENTION OF ALGORITHMIC DISCRIMINATION

Protecting from Discrimination and
other Harms Caused by Algorithms
through Privacy in the EU: Possibilities,
Shortcomings and Proposals

ALBA SORIANO ARNANZ

DATA PROTECTION FOR THE PREVENTION OF ALGORITHMIC DISCRIMINATION

Protecting from Discrimination
and other Harms Caused by
Algorithms through Privacy
in the EU: Possibilities,
Shortcomings and Proposals

R.218250

THOMSON REUTERS

ARANZADI



Summary

Página

ACKNOWLEDGEMENTS	19
PREFACE	21
INTRODUCTION	25

CHAPTER I

AN INTRODUCTION TO ALGORITHMIC DECISION- MAKING	33
--	----

1. Big data	33
1.1. <i>The three Vs in big data</i>	34
1.2. <i>The need to process (raw) big data</i>	35
1.3. <i>The fourth V: value</i>	36
2. Data processing tools and technologies	38
2.1. <i>Machine learning and data mining</i>	38
2.2. <i>Supervised and unsupervised learning</i>	40
2.3. <i>Algorithms and models</i>	42
3. The application of automated systems	44
3.1. <i>The use of algorithms by the private sector</i>	46
3.1.1. <i>Scoring individuals</i>	46
3.1.1.1. <i>The banking sector and the expansion of credit scores</i>	47
3.1.1.2. <i>Healthcare</i>	49
3.1.1.3. <i>Human resources</i>	50
3.1.2. <i>Consumer profiling and advertising</i>	51
3.2. <i>The use of algorithms by the public sector</i>	52

3.2.1.	The use of algorithms in public service management and provision	52
3.2.1.1.	The use of algorithms in public aid and welfare programmes	54
3.2.1.2.	Automation of public services, aid and welfare programmes and the perpetuation of inequality	55
3.2.2.	The use of algorithms in public administration's regulatory and coercive activity: law enforcement	57
3.2.2.1.	Police departments and the criminal justice system	57
3.2.2.2.	Other algorithmic applications in the exercise of administrative regulatory and coercive powers	59
4.	Types of algorithmic decision-making	60
4.1.	<i>Automatic and autonomous systems</i>	60
4.2.	<i>Automated and semi-automated systems</i>	61
4.3.	<i>Profiling and automated decision-making: descriptive, predictive, classification and recommendation purposes</i>	62

CHAPTER II

	RISKS AND HARMS GENERATED BY THE USE OF AUTOMATED SYSTEMS	65
1.	Problems and risks for the protection of the rights of individuals subjected to automated decision-making	65
1.1.	<i>Biases and errors</i>	66
1.1.1.	Biased humans and accurate machines	66
1.1.2.	Measuring baseball vs. measuring humans	67
1.1.3.	Human bias and machine error	68
1.1.4.	The technological heuristic	70
1.2.	<i>Algorithmic discrimination</i>	71
1.3.	<i>Risks to dignity: individuality, autonomy and privacy</i>	74

1.4. <i>Transparency, due process and traceability</i>	77
1.4.1. <i>Transparency</i>	77
1.4.2. <i>Justification and understandability</i>	79
1.4.3. <i>Participation and due process</i>	80
1.4.4. <i>Traceability</i>	82
2. The legitimacy and legality of public automated decision-making	83
2.1. <i>Transparency and justification of public decisions</i>	84
2.1.1. <i>The private exercise of inherently public tasks</i>	85
3. Market failures and intervention in the private sector	87
3.1.1. <i>The precautionary principle</i>	88
3.1.2. <i>Market failures and other problems generated by the data services sector</i>	92
3.1.2.1. <i>Negative externalities</i>	92
3.1.2.2. <i>Monopolistic behaviour</i>	93
3.1.2.3. <i>Asymmetric information, imperfect rationality and transaction costs</i>	96

CHAPTER III

THE INFORMATIONAL PRIVACY FRAMEWORK. GENERAL ASPECTS	99
1. The informational privacy framework as a solution for the harms caused by algorithms	99
1.1. <i>The right to data protection as an anti-classification instrument</i>	102
1.2. <i>The fundamental right to data protection</i>	106
2. EU and US privacy traditions	107
3. The scope of application of informational privacy regulations	111
3.1. <i>Anonymisation</i>	112
3.2. <i>Pseudonymisation</i>	114
3.3. <i>Scope of application of the EU's data protection framework</i>	115

4. Privacy principles	116
4.1. <i>Data processing principles: lawfulness, fairness, transparency, integrity and confidentiality</i>	<i>117</i>
4.2. <i>Data collection principle: purpose limitation</i>	<i>120</i>
4.3. <i>Data and storage requirements: data minimisation, accuracy and storage limitation</i>	<i>121</i>

CHAPTER IV

PROHIBITIONS TO ACCESS AND PROCESS INFORMATION	123
---	------------

1. The US approach to protection through data collection and processing prohibitions	124
1.1. <i>The Health Insurance Portability and Accountability Act</i>	<i>124</i>
1.2. <i>The Americans with Disabilities Act</i>	<i>126</i>
1.3. <i>The Genetic Information Nondiscrimination Act</i>	<i>127</i>
1.4. <i>The Family Educational Rights and Privacy Act (FERPA)</i>	<i>127</i>
1.5. <i>The Fair Credit Reporting Act (FCRA) and Equal Credit Opportunity Act (ECOA)</i>	<i>128</i>
2. Privacy as anti-discrimination through general prohibitions in the GDPR	129
2.1. <i>Processing special categories of personal data</i>	<i>129</i>
2.1.1. <i>Scope of the prohibition</i>	<i>131</i>
2.1.1.1. <i>Personal scope of application: search engine operators</i>	<i>131</i>
2.1.1.2. <i>Material scope of application: the proxy problem</i>	<i>132</i>
2.1.1.3. <i>Solutions for the discrimination by proxy problem</i>	<i>133</i>
2.1.2. <i>Processing of personal data relating to criminal convictions and offences</i>	<i>136</i>
2.2. <i>The right (or general prohibition) not to be subject to decisions based solely on automated processing, including profiling</i>	<i>136</i>

2.2.1.	The right not to be subject to a decision based solely on automated processing, including profiling	136
2.2.2.	Exceptions to the right (prohibition) recognised in article 22 and safeguards	137
2.2.3.	Special protections for decisions based solely on the automated processing of special categories of personal data	139
2.2.4.	Issues raised with regard to the scope of article 22.1	140
2.2.4.1.	Decisions based solely on automated processing	140
2.2.4.2.	Legal or significantly similar effects	142
2.2.5.	Analysis of the exceptions to the right not to be subject to a decision based solely on automated processing, including profiling	144
2.2.5.1.	Necessary for entering into, or performance of, a contract	144
2.2.5.2.	Authorised by EU or member state law	144
2.2.5.3.	The data subject's explicit consent	145
2.2.5.4.	Additional elements that must concur for applying the exceptions to the processing of special categories of personal data	146
3.	Prohibitions in the Directive for personal data protection in law enforcement and the criminal justice system	148
3.1.	<i>Harmonisation and scope of application</i>	<i>149</i>
3.2.	<i>Processing special categories of personal data within the scope of Directive 2016/680</i>	<i>150</i>
3.3.	<i>The prohibition of decisions based solely on automated processing, including profiling</i>	<i>151</i>
4.	Shortcomings in the prohibitions contained in the eu personal data protection framework	155

CHAPTER V

TECHNOLOGICAL DUE PROCESS RIGHTS	157
1. The informational self-determination approach	157
2. Transparency: The rights to information, access and explanation	160
2.1. <i>Information, access and explanation rights in the GDPR</i>	161
2.1.1. The right to be informed	162
2.1.1.1. The intended purposes of the processing	162
2.1.1.2. Meaningful information about the logic involved, significance and envisaged consequences	164
2.1.2. The right to access	165
2.1.3. The right to explanation	166
2.1.3.1. Internal limits to the right to explanation	167
2.1.3.2. External limits to the right to explanation	168
i) The conflict with trade secrets and intellectual property	168
ii) State secrets and public interests	170
2.1.3.3. How the right to explanation can be made effective	171
2.2. <i>Information, access and explanation rights in Directive 2016/680 for data protection in law enforcement: the conflict with state and public security</i>	172
2.3. <i>Information, access and explanation rights in US regulatory instruments</i>	174
2.3.1. The Fair Credit Reporting Act (FCRA)	174
2.3.2. The California Consumer Privacy Act (CCPA) and California Privacy Rights Act (CPRA)	175
2.4. <i>A few final remarks with regard to the transparency principle and the rights that derive from it</i>	176

3. The right to be heard and contest decisions: the right to an effective remedy	178
3.1. <i>The right to be heard and contest decisions in the GDPR</i>	178
3.1.1. Data subjects' due process rights in art. 22	178
3.1.1.1. The right to obtain human intervention	179
3.1.1.2. The right to express his or her point of view	179
3.1.1.3. The right to challenge the decision	181
3.1.2. Individual rights to be heard and challenge decisions recognised outside of article 22 of the GDPR	181
3.1.2.1. The rights to data portability, rectification, erasure and restriction of processing	181
3.1.2.2. The right to object	184
3.1.2.3. The rights to lodge complaints before supervisory authorities and to judicial remedies	185
3.2. <i>The rights to be heard and challenge decisions in Directive 680/2016</i>	185
3.3. <i>Rights to be heard and challenge decisions in US regulatory instruments</i>	186
3.3.1. The Fair Credit Reporting Act	186
3.3.2. The California Consumer Privacy Act and California Privacy Rights Act	186
4. The need to complement technological due process rights with a system for algorithmic oversight and control	187

CHAPTER VI

REGULATORY MECHANISMS FOR SYSTEM TRANSPARENCY AND ACCOUNTABILITY THROUGH DATA PROTECTION	189
1. Regulatory frameworks	189
1.1. <i>Self-regulation</i>	189
1.2. <i>Co-regulation (or regulated self-regulation)</i>	191
1.3. <i>Regulation (state intervention)</i>	192

	<i>Página</i>
2. The GDPR as a system of governance	193
3. System transparency and accountability	194
4. Regulatory Tools For System Transparency And Accountability	196
4.1. Rule-setting mechanisms	196
4.1.1. Safe harbour and privacy shields	196
4.1.2. Codes of conduct	196
4.1.3. Technical and organisational standards	199
4.2. Control mechanisms	203
4.2.1. Certification mechanisms	203
4.2.1.1. General issues	203
4.2.1.2. Certification in the GDPR	206
4.2.2. Data protection impact assessments	207
4.2.3. Re-certification, DPIA reviews and audits	210
4.3. Enforceability mechanisms	211
4.3.1. Ethics committees and data protection officers	211
4.3.2. The European Data Protection Board and Data Protection Authorities	213
4.3.3. Penalties	214

CHAPTER VII

THE PRIVACY FRAMEWORK: SHORTCOMINGS AND TENSIONS	215
1. General shortcomings of the privacy approach	216
1.1. The unrealistic expectations of anonymisation	216
1.2. The limits of personal data protection	217
1.2.1. Group profiling	217
1.2.2. Output data	217
1.2.3. Failure to focus on varieties of processing	218
2. The shortcomings of the informational-self determination approach	218
2.1. The myth of consent and the privacy paradox	219

2.2.	<i>Asymmetric information and burdens</i>	221
2.3.	<i>Creating systemic inaccuracies</i>	223
2.4.	<i>The difficulty of detecting systemic errors</i>	223
3.	Privacy approaches are not appropriate for the use of algorithms by the public sector	225
3.1.	<i>Private sector limits to transparency for algorithms used by public bodies</i>	226
3.1.1.	<i>Intellectual property and the Spanish “energy social bond”</i>	226
3.1.2.	<i>Administrative courts granting transparency</i>	229
3.2.	<i>Banning the use of algorithms in the public sector: the Dutch “SyRI” case</i>	230
3.3.	<i>The regulatory nature of algorithms employed by public administrations</i>	232
3.3.1.	<i>Algorithms used by public administrations are legal instruments</i>	233
3.3.2.	<i>Algorithms are regulatory instruments</i>	234
3.3.2.1.	<i>Proposals that reject the regulatory nature of algorithms</i>	235
3.3.2.2.	<i>Administrative court of Lazio-Roma, Judgment No. 3769</i>	238
3.3.2.3.	<i>Solely automated non-binding and semi-automated decision making</i>	239
3.3.2.4.	<i>The importance of recognising the regulatory nature of algorithms</i>	239
3.3.3.	<i>The principle of legality must apply to the public use of algorithms</i>	240
3.3.4.	<i>Frictions between traditional and algorithmic regulation</i>	241
4.	The shortcomings of accountability mechanisms	242
5.	The relationship between personal data protection, equality and non-discrimination	246
5.1.	<i>The privacy vs. antidiscrimination dilemma</i>	246
5.1.1.	<i>Less information can lead to wrong inferences</i>	247

5.1.2. Anti-classification does not prevent indirect algorithmic discrimination	249
5.1.3. Anti-classification through privacy does not solve group disadvantage and can reinforce it	250
5.2. <i>Combining the anti-discrimination and data protection frameworks</i>	252

CHAPTER VIII

POSSIBILITIES AND PROPOSALS FOR THE REGULATION OF ALGORITHMS	255
1. Trade-offs in the regulation of algorithms	256
2. The need for more algorithmic transparency	259
3. A system of public intervention to control algorithms	264
3.1. <i>Organisational options</i>	264
3.1.1. Algorithmic control mainstreaming	265
3.1.2. Creating a non-independent supervisory task force or body	266
3.1.3. An independent supervisory agency	266
3.2. <i>Risk-based market approval of algorithms</i>	268
3.2.1. The three (plus two) tier system	270
3.2.1.1. Prohibited algorithmic systems	270
3.2.1.2. High-risk algorithmic systems	271
i) Administrative testing, documentation and general explanation requirements	273
ii) Justification and explainability requirements	274
iii) The proportionality analysis of pre-market authorisations	277
iv) Specific requirements for public sector algorithms included in this category	278

SUMMARY

	<i>Página</i>
3.2.1.3. Medium-risk algorithmic systems	281
3.2.1.4. Low-risk algorithmic systems	281
3.2.1.5. Non-risky algorithmic systems	282
3.2.2. System enforcement	282
3.3. <i>Public procurement as a mechanism to prevent the risks of the public and private use of algorithms</i>	283
3.4. <i>Establishing a “best available techniques” regime</i>	283
3.5. <i>Using algorithms to detect discrimination</i>	284
3.6. <i>Empowering individuals through understandable information: choice architectures</i>	285
3.7. <i>Increased communication between disciplines and establishing general principles upon which to construct automated systems</i>	285
CONCLUSIONS	287
BIBLIOGRAPHY	289

Thomson Reuters ProView. Guía de uso