



THREE LIABILITY  
REGIMES *for*  
ARTIFICIAL INTELLIGENCE

*Algorithmic Actants, Hybrids, Crowds*

ANNA BECKERS AND GUNTHER TEUBNER

## THREE LIABILITY REGIMES FOR ARTIFICIAL INTELLIGENCE

This book proposes three liability regimes to combat the wide responsibility gap caused by AI systems – vicarious liability for autonomous software agents (actants); enterprise liability for inseparable human-AI interactions (hybrids); and collective fund liability for interconnected AI systems (crowds).

Based on information technology studies, the book first develops a threefold typology that distinguishes individual, hybrid and collective machine behaviour. A subsequent social sciences analysis specifies the socio-technical configurations of this threefold typology and theorises their social risks when being used in social practices: actants raise the risk of digital autonomy, hybrids the risk of double contingency, crowds the risk of opaque interconnections. The book demonstrates that it is these specific risks to which the law needs to respond, by recognising personified algorithms as vicarious agents, human-machine associations as collective enterprises, and interconnected systems as risk pools – and by developing corresponding liability rules.

The book relies on a unique combination of information technology studies, sociological configuration and risk analysis, and comparative law. This unique approach uncovers recursive relations between types of machine behaviour, emergent socio-technical configurations, their concomitant risks, the legal conditions of liability rules, and the ascription of legal status to the algorithms involved.



# Three Liability Regimes for Artificial Intelligence

*Algorithmic Actants, Hybrids, Crowds*

Anna Beckers  
and  
Gunther Teubner

• H A R T •

OXFORD • LONDON • NEW YORK • NEW DELHI • SYDNEY

HART PUBLISHING  
Bloomsbury Publishing Plc  
Kemp House, Chawley Park, Cumnor Hill, Oxford, OX2 9PH, UK  
1385 Broadway, New York, NY 10018, USA  
29 Earlsfort Terrace, Dublin 2, Ireland

HART PUBLISHING, the Hart/Stag logo, BLOOMSBURY and the Diana logo are  
trademarks of Bloomsbury Publishing Plc

First published in Great Britain 2021

Copyright © Anna Beckers and Gunther Teubner, 2021

Anna Beckers and Gunther Teubner have asserted their right under the Copyright, Designs and  
Patents Act 1988 to be identified as Authors of this work.

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any  
means, electronic or mechanical, including photocopying, recording, or any information storage  
or retrieval system, without prior permission in writing from the publishers.

While every care has been taken to ensure the accuracy of this work, no responsibility for  
loss or damage occasioned to any person acting or refraining from action as a result of any  
statement in it can be accepted by the authors, editors or publishers.

All UK Government legislation and other public sector information used in the work is  
Crown Copyright ©. All House of Lords and House of Commons information used in  
the work is Parliamentary Copyright ©. This information is reused under the terms  
of the Open Government Licence v3.0 ([http://www.nationalarchives.gov.uk/doc/  
open-government-licence/version/3](http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3)) except where otherwise stated.

All Eur-lex material used in the work is © European Union, <http://eur-lex.europa.eu/>, 1998–2021.

A catalogue record for this book is available from the British Library.

Library of Congress Cataloging-in-Publication data

Names: Beckers, Anna, 1982- author. | Teubner, Gunther, author.

Title: Three liability regimes for artificial intelligence : algorithmic actants, hybrids,  
crowds / Anna Beckers and Gunther Teubner.

Description: Oxford ; New York : Hart, 2021. | Includes bibliographical references and index.

Identifiers: LCCN 2021044688 (print) | LCCN 2021044689 (ebook) |  
ISBN 9781509949335 (hardback) | ISBN 9781509949373 (paperback) |  
ISBN 9781509949359 (pdf) | ISBN 9781509949342 (Epub)

Subjects: LCSH: Artificial intelligence—Law and legislation. | Liability (Law)

Classification: LCC K564.C6 B43 2021 (print) | LCC K564.C6 (ebook) |  
DDC 343.09/99—dc23/eng/20211028

LC record available at <https://lcn.loc.gov/2021044688>

LC ebook record available at <https://lcn.loc.gov/2021044689>

ISBN: HB: 978-1-50994-933-5  
ePDF: 978-1-50994-935-9  
ePub: 978-1-50994-934-2

Typeset by Compuscript Ltd, Shannon

To find out more about our authors and books visit [www.hartpublishing.co.uk](http://www.hartpublishing.co.uk). Here you will find  
extracts, author information, details of forthcoming events and the option to sign up for our newsletters.

---

## PREFACE

---

In this book, we propose three liability regimes for addressing the considerable responsibility gaps caused by AI-systems: Vicarious liability for autonomous software agents (actants), enterprise liability for inseparable human-AI interactions (hybrids) and collective fund liability for interconnected AI systems (crowds). The liability regimes serve as finely tuned reactions to liability gaps of different quality. Instead of overgeneralising a one-size-fits-all liability or undergeneralising a sectorally fragmented liability along with the various contexts in which AI is used, we focus on three fundamental risks that AI systems pose: autonomous decision-making, association with humans, and systemic interconnectivity.

Methodologically, our book suggests new interdisciplinary ways of thinking of the interrelation between technology and liability law. In contrast to the regularly observed short-cut that translates technological properties directly into liability rules, we place the emphasis on the social sciences as an intermediary discipline between AI technology and law. The social sciences help identify the social-technical configurations in which AI systems appear and theorise their social risks that law needs to respond to within its own system of rules. We propose to introduce the concept of ‘socio-digital institutions’. Algorithms do not have as such the ontological qualities of an actor that allow them to engage in social relations and communicate with humans. Only once algorithms are part of socio-digital institutions, these institutions will, according to their normative premises, obtain communicative capacities and qualify as actors. Our approach also differs from the typical focus that lawyers place on economics and thus the costs and benefits of liability systems. Instead, we integrate insights from social theory, moral philosophy, and the philosophy of technology. These insights are particularly helpful for dealing with complex issues such as personification of algorithms, emergent properties of human-algorithm associations and distributed cognition of interconnected networks.

We recognise that liability rules remain, to a large extent, fragmented along national lines. Therefore, our legal analysis contains a comparative dimension. To provide a solid basis for algorithms’ status in law, we focus on the current discussion in the civil law world with a particular view to the specifics of German law, and in the common law world, particularly in the US and English law. Whenever relevant, we also integrate the European dimension of the topic. Our comparative analysis follows a method that Collins has coined

‘comparative sociological jurisprudence’.<sup>1</sup> Sociological jurisprudence analyses socio-digital institutions and their inherent risks to framing the relevant legal categories; comparative sociological jurisprudence uses this analysis with a view to different legal systems and the specifics of national doctrines. Our analysis of the various risks attempts to identify the most suitable legal categories for handling this problem. In spelling out how these categories are applied, the study then accounts for liability laws in national legal orders, their concepts in legal doctrine, and their basic principles.

Combining interdisciplinary analysis on socio-digital institutions and comparative legal dogmatics of liability law provides a path on how the law can respond to the real and pressing current liability gaps. At the same time, it is to be read as a proposal for a general way of thinking about the future of liability law in an era of technological advancement and related social risks.

The book has benefitted from intense discussions with many colleagues. Our thanks go especially to Marc Amstutz, Alfons Bora, Carmela Camardi, Ricardo Campos, Elena Esposito, Pasquale Femia, Andreas Fischer-Lescano, Malte Gruber, Albert Ingold, Günter Küppers, Dimitrios Linardatos, Martin Schmidt-Kessel, Juliano Maranhão, Marc Mölders, Michael Monterossi, Daniel On, Oren Perez, Valentin Rauer, Jan-Erik Schirmer, Thomas Vesting, Gerhard Wagner, Dan Wielsch, and Rudolf Wiethölter. We also like to thank the three anonymous reviewers for their careful reading and commenting on the proposal and manuscript. Anna Huber and Dirk Hildebrandt have provided substantial historical art expertise on Max Ernst and the overpainting *figure ambigue* that we chose as the image for the cover. We also thank the team at Hart Publishing, most notably Roberta Bassi and Rosemarie Mearns, for sharing our enthusiasm for this book idea and for their professional guidance in the book’s production.

Anna Beckers & Gunther Teubner  
July 2021

<sup>1</sup> Fundamentally, H Collins, *Introduction to Networks as Connected Contracts* (Oxford, Hart, 2011) 25 ff. See also for an extensive use of this method A Beckers, *Enforcing Corporate Social Responsibility Codes: On Global Self-Regulation and National Private Law* (Oxford, Hart, 2015) chs 2 and 6.



---

# CONTENTS

---

<i>Preface</i> .....	v
<b>1. Digitalisation: The Responsibility Gap</b> .....	<b>1</b>
I. The Problem: The Dangerous <i>Homo Ex Machina</i> .....	1
A. Growing Liability Gaps.....	3
B. Scenarios.....	4
C. Current Law's Denial of Reality.....	6
II. The Overshooting Reaction: Full Legal Subjectivity for E-Persons? .....	8
III. Our Solution: Differential Legal Status Ascriptions for Algorithms.....	10
A. Algorithms in Social and Economic Contexts .....	10
B. (Legal) Form Follows (Social) Function.....	12
IV. Our Approach: Three Digital Risks.....	14
A. 'Socio-Digital Institutions' as Intermediaries between Technology and Law.....	14
B. A Typology of Machine Behaviour .....	17
C. A Typology of Socio-Digital Institutions .....	18
D. A Typology of Liability Risks.....	20
<b>2. Autonomy and Personification</b> .....	<b>23</b>
I. Artificial Intelligence as Actants.....	23
A. Anthropomorphism? .....	23
B. Actants and Action Attribution.....	25
C. Communication with Actants .....	27
II. Gradualised Digital Autonomy.....	31
A. Social Attribution of Autonomy .....	31
B. Legal Criteria of Autonomy .....	34
C. Our Solution: Decision under Uncertainty.....	37
III. Autonomy and Legal Personhood.....	41
A. Against Personification? .....	41
B. Uniform Personification? .....	43
C. Socio-Digital Institutions and Legal Status.....	43
<b>3. Actants: Autonomy Risk</b> .....	<b>45</b>
I. Socio-Digital Institution: Digital Assistance.....	45
II. The Autonomy Risk.....	46



III.	Algorithmic Contract Formation .....	49
A.	Invalidity of Algorithmic Contracts? .....	49
B.	Algorithms as Mere Tools? .....	51
C.	Our Solution: Agency Law for Electronic Agents .....	54
D.	Limited Legal Personhood – Constellation One .....	55
E.	The Digital Agent’s Legal Declaration .....	57
F.	Digital Assistance and the Principal-Agent Relation .....	60
G.	Overstepping of Authority? .....	62
IV.	Contractual Liability .....	64
A.	The Dilemma of the Tool-Solution .....	65
B.	Our Solution: Vicarious Performance .....	67
C.	Limited Legal Personhood – Constellation Two .....	68
V.	Non-Contractual Liability .....	71
A.	Fault-Based Liability? .....	72
B.	Product Liability? .....	74
C.	Strict Causal Liability for Dangerous Objects and Activities? .....	76
D.	Our Solution: Vicarious Liability in Tort .....	79
E.	Limited Legal Personhood – Constellation Three .....	84
F.	The ‘Reasonable Algorithm’ .....	84
G.	Who is Liable? .....	86
<b>4.</b>	<b>Hybrids: Association Risk .....</b>	<b>89</b>
I.	Socio-Digital Institution: Human-Machine Associations .....	90
A.	Emergent Properties .....	90
B.	Hybridity .....	92
C.	The Organisational Analogy .....	94
II.	The Association Risk .....	96
III.	Solution <i>de lege ferenda</i> : Hybrids as Legal Entities? .....	97
IV.	Our Solution <i>de lege lata</i> : Enterprise Liability for Human-Machine Networks .....	99
A.	Human-Machine Interactions as Networks .....	99
B.	Networks and Enterprise Liability .....	101
C.	Action Attribution and Liability Attribution in Hybrids .....	103
D.	Liable Actors .....	103
E.	Pro-Rata Network Share Liability .....	106
F.	External Liability Concentration: ‘One-Stop-Shop’ Approach .....	108
G.	Internal Liability Distribution: Pro Rata Network Share .....	109
V.	Conclusion .....	110
<b>5.</b>	<b>Multi-Agent Crowds: Interconnectivity Risk .....</b>	<b>111</b>
I.	Socio-Digital Institution: Exposure to Interconnectivity .....	111
A.	Non-Communicative Contacts .....	112
B.	Distributed Cognitive Processes .....	114
II.	The Interconnectivity Risk .....	116

III.	Mismatch of New Risks and Existing Solutions .....	120
A.	Applying Existing Categories.....	120
B.	Vicarious or Product Liability of the Whole Interconnected AI-System? .....	121
C.	Collective Liability of Actors Connected to the ‘Technical Unit’? .....	122
IV.	Our Solution: Socialising the Interconnectivity Risk .....	125
A.	Entire or Partial Socialisation? .....	125
B.	Risk Pools Decried by Law .....	126
C.	Public Administration of the Risk Pool: The Fund Solution .....	128
D.	Financing: Ex-Ante and Ex-Post Components .....	129
E.	Participation and Administration .....	130
F.	Compensation and Recovery Action .....	133
G.	Global Interconnectivity and National Administration.....	135
<b>6.</b>	<b>Conclusion: Three Liability Regimes and Their Interrelations.....</b>	<b>138</b>
I.	Synopsis and Rules .....	138
A.	Digital Assistance: Vicarious Liability.....	139
B.	Human-Machine Associations: Enterprise Liability.....	139
C.	Interconnectivity: Fund Liability.....	139
II.	Socio-Digital Institutions and Liability Law.....	140
A.	One-Size-Fits-All or Sector-Specific Piecemeal Approach? .....	140
B.	Socio-Digital Institutions .....	142
C.	Institution-Generating Liability Law .....	145
D.	Differential Treatment of Liable Actors.....	149
E.	Calibrating Legal Personhood for AI.....	152
III.	Interactions between Three Liability Regimes.....	153
A.	Criteria for Delineating the Applicable Liability Regime.....	153
B.	No Priority for a Liability Regime .....	155
IV.	Exemplary Cases.....	158
A.	Vicarious Liability: Robo-Advice.....	158
B.	Enterprise Liability: Hybrid Journalism.....	159
C.	Collective Funds: Flash Crash .....	161
D.	Finale: Google Autocomplete .....	163
	<i>Bibliography</i> .....	167
	<i>Index</i> .....	183



---

## Digitalisation: The Responsibility Gap

---

### I. The Problem: The Dangerous *Homo Ex Machina*

'*Figure ambiguë*' – the overpainting, which is reproduced on the cover of this book, was produced by Max Ernst, one of the protagonists of dadaism/surrealism. In 1919, he already expressed his unease with the excessive ambivalences of modern technology. His work is simultaneously celebratory about the dynamism and energy of the machine utopia and sarcastic about its dehumanising consequences. On the painting's right side, Ernst creates a serene joyful atmosphere that seems to symbolise the ingenious inventions of modern science. Mechanically animated letters of the alphabet are connected to each other in complex arrangements and seem to be transformed into strange machines. Via metamorphosis or double identity, these non-human figures appear to substitute human bodies; they jump, dance, and even fly. These *homines ex machina* 'carry off a triumph of mobility: through rotation, doubling, shifting, reflection, and optical illusion'.<sup>1</sup>

Abruptly, the atmosphere changes on the painting's left side. The symbols change their colour, become dark, appear to be brutal and threatening. In the upper left corner, a black sun, which is again made up of strange symbols forming a sinister face, is throwing its dark light over the world. With this painting and many others, Max Ernst expressed his ambivalent attitude toward the logic, rationality and aesthetics of the modern perfect machine world, which had the potential to turn into absurdity, irrationality and brutality.<sup>2</sup> Ernst 'was looking for ways to register social mechanisms and truths as well as to symbolise with artistic techniques their more profound structure. Probably, it is an attempt to grasp a social subconscious in the historical moment when the totalitarian potential of technology became imaginable'.<sup>3</sup>

Today, Max Ernst's surrealistic dream seems to become the new reality. Algorithms are the emblematic *figures ambiguës* of our time, which even radicalise

<sup>1</sup> R Uhl, *Prehistoric Future: Max Ernst and the Return of Painting Between Wars* (Chicago, Chicago University Press, 2004) 26, 28.

<sup>2</sup> V Becchetti, 'Max Ernst: Il surrealista psicoanalitico', (2020) *LoSpessore – Opinioni, Cultura e Analisi della Società* [www.lospessore.com/10/11/2020/max-ernst-il-surrealista-psicoanalitico/](http://www.lospessore.com/10/11/2020/max-ernst-il-surrealista-psicoanalitico/); E Adamowicz, *Dada Bodies: Between Battlefield and Fairground* (Manchester, Manchester University Press, 2019) chs 4 and 8.

<sup>3</sup> This is how the art historian Anna Huber interpreted Max Ernst's work in a letter to the authors.

## 2 Digitalisation: The Responsibility Gap

the ambivalence of machine automatons by an enigmatic ‘artificial intelligence’. Like the alphabetic letters in Max Ernst’s painting, algorithms, at first sight, are nothing but innocent chains of symbols. In their electronic metamorphosis, these symbols begin to live, jump, dance, fly. What is more, they bring into existence a new world of meaning. Their *creatio ex nihilo* promises a better future for mankind. Big data and algorithmic creativity symbolise the hopes of expanding or substituting the cognitive capacities of the human mind. But this is only the bright side of their excessive ambivalence. There is a threatening dark side to the brave new world of algorithms, who, after the first phase of enthusiasm, are now often perceived as nightmarish monsters. ‘Perverse instantiation’ results when intelligent machines run out of human control: the individual algorithm efficiently satisfies the goal set by the human participant but chooses a means that violates the human’s intentions.<sup>4</sup> Moreover, a strange hybridity emerges when humans and machines begin not only to communicate but also to create supervenient *figures ambiguas* with undreamt-of potentially damaging characteristics. And, the most threatening situation arises, as symbolised in Max Ernst’s dark sun, in the dangerous exposure of human beings to an opaque algorithmic environment that remains uncontrollable.

How does contemporary law deal with algorithmic *figures ambiguas*? That is the theme of this book, exemplified by the law of liability for algorithmic failures. Law mirrors the excessive ambivalence of the world of algorithms. On their bright side, law welcomes algorithms as powerful instruments in the service of human needs. Law opens itself to algorithms, conferring to them even a quasi-magic *potestas vicaria* so that they can participate as autonomous agents in transactions on the market. However, on their dark side, current law reveals remarkable deficiencies. Liability law is not at all prepared to counteract the algorithms’ new dangers. Ignoring the potential threats stemming from their autonomy, the law treats algorithms not any different from other tools, machines, objects, or products. If they create damages, current product liability is supposed to be the appropriate reaction.

But that is too easy. Compared to familiar situations of product liability, with the arrival of algorithms, ‘the array of potential harms widens, as to the product is added a new facet – intelligence.’<sup>5</sup> The *figures ambiguas* that invade private law territories are not simply hazardous objects but uncontrollable subjects – robots, software agents, cyborgs, hybrids, computer networks – some with a high level of autonomy and the ability to learn. With their restless energy, they generate new kinds of undreamt-of hazards for humans and society.

<sup>4</sup> N Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford, Oxford University Press, 2017) 146 ff.

<sup>5</sup> O Rachum-Twaig, ‘Whose Robot is it Anyway? Liability for Artificial-Intelligence-Based Robots’, [2020] *University of Illinois Law Review* 1141, 1149.

In the legal debate, defensive arguments abound to keep these alien species at a distance. The predominant position in legal scholarship argues with astonishing self-confidence that the rules on contract formation and liability in contract, tort and product liability are, in their current form, well equipped to deal with the hazards of such new digital species. According to this opinion, there is no need of deviating from the established methods of action and liability attribution. Computer behaviour is nothing but behaviour of the humans behind the machine. Autonomous AI systems are legally treated, so the argument goes, without problems as mere machines, as human tools, as willing instruments in the hands of their human masters.<sup>6</sup>

## A. Growing Liability Gaps

However, private law categories cannot avoid responding to the current and very real problems that algorithms cause when acquiring autonomy.<sup>7</sup> A new phenomenon called ‘active digital agency’ is causing the problems:

The more autonomous robots will become, the less they can be considered as mere tools in the hand of humans, and the more they obtain active digital agency. In this context, issues of responsibility and liability for behaviour and possible damages resulting from the behaviour would become pertinent.<sup>8</sup>

Unacceptable gaps in responsibility and liability – this is why private law needs to change its categories fundamentally. Given the rapid digital developments, the gaps have already opened today.<sup>9</sup> Software agents and other AI systems inevitably cause these gaps because their actions are unpredictable and thus entail a

<sup>6</sup>For US-Law: Restatement (Third) of Agency Law § 1.04 cmt. e. (2006); A Bertolini, ‘Robots as Products: The Case for a Realistic Analysis of Robot Applications and Liability Rules’, (2013) 5 *Law, Innovation & Technology* 214. For English law: *Software Solutions Partners Ltd. v HM Customs & Excise* [2007] EWHC Admin 971, para 67. For German law: FJ Säcker et al., *Münchener Kommentar zum Bürgerlichen Gesetzbuch. Band 1* 8th edn (Munich, C.H. Beck, 2018), Introduction to § 145, 38 (Busche).

<sup>7</sup>Here, we refer to digital autonomy in a rather loose sense. Later on, we will discuss extensively its precise meaning, particularly in ch 2, II.

<sup>8</sup>N van Dijk, ‘In the Hall of Masks: Contrasting Modes of Personification’, in M Hildebrandt and K O’hara (eds), *Life and the Law in the Era of Data-Driven Agency* (Cheltenham, Edward Elgar, 2020) 231. The concept ‘active digital agency’ has been introduced by R Clarke, ‘The Digital Persona and its Application to Data Surveillance’, (1994) 10 *Information Society* 77.

<sup>9</sup>The responsibility gaps have alarmed the European Parliament resulting in the Resolution of 16 February 2017 with Recommendations to the Commission on Civil Law Rules on Robotics, 2015/2103(INL) para 10; European Parliament, Civil Liability Regime for Artificial Intelligence, Resolution of 20 October 2020, 2020/1012(INL), paras 49–59. They also informed the EU Commission’s understanding on liability for AI: European Commission, ‘Report on the Safety and Liability Implications of Artificial Intelligence, The Internet of Things and Robotics’, COM(2020) 64 final, 16 (with particular view to gaps in product liability). On the novel liability risk of digital autonomy, see, eg: S Dyrkolbotn, ‘A Typology of Liability Rules for Robot Harms’, in M Aldinhas Ferreira et al. (eds), *A World with Robots: Intelligent Systems, Control and Automation* (Cham, Springer, 2017) 121 f.

#### 4 *Digitalisation: The Responsibility Gap*

massive loss of control for human actors. At the same time, society is becoming increasingly dependent on autonomous algorithms on a large scale, and it is improbable that society will abandon their use.<sup>10</sup>

Of course, lawyers' resistance to granting algorithms the status of legal capacity or even personhood is understandable. After all, '[t]he fact is, that each time there is a movement to confer rights onto some new "entity", the proposal is bound to sound odd or frightening or laughable.'<sup>11</sup> But despite the oddity of 'algorithmic persons', the growing responsibility gaps confront private law with a radical choice: either it assigns AI-systems an independent legal status as responsible actors or accepts an increasing number of accidents without anyone being responsible for them. The dynamics of digitalisation are constantly creating responsible-free spaces that will expand in the future.<sup>12</sup>

### B. Scenarios

When using the serious threat of increasing liability gaps, it is of course crucial to clearly identify such gaps in the first place. Information science describes typical responsibility gaps in the following scenarios: Deficiencies arise in practice when the software is produced by teams, when management decisions are just as important as programming decisions, when documentation of requirements and specifications plays a significant role in the resulting code, when, despite testing code accuracy, a lot depends on 'off-the-shelf' components whose origin and accuracy are unclear, when the performance of the software is the result of the accompanying checks and not of program creation, when automated instruments are used in the design of the software, when the operation of the algorithms is influenced by its *interfaces* or even by system traffic, when the software interacts in an unpredictable manner, or when the software works with probabilities or is adaptable or is the result of another program.<sup>13</sup>

These scenarios produce the most critical liability gaps that the law has so far encountered.<sup>14</sup>

#### *i. Machine Connectivities*

The most challenging liability gap arises in multiple agent systems when several computers are closely interconnected in an algorithmic network and create

<sup>10</sup> A Matthias, *Automaten als Träger von Rechten* 2nd edn (Berlin, Logos, 2010) 15.

<sup>11</sup> CD Stone, *Should Trees Have Standing? Toward Legal Rights for Natural Objects* (Los Altos, Kaufmann, 1974) 8.

<sup>12</sup> This is the central and well-documented thesis of Matthias, *Automaten* 111.

<sup>13</sup> L Floridi and JW Sanders, 'On the Morality of Artificial Agents', in M Anderson and SL Anderson (eds), *Machine Ethics* (Cambridge, Cambridge University Press, 2011) 205.

<sup>14</sup> For a detailed list of liability gaps for wrongful acts of algorithms, see M Bashayreh et al., 'Artificial Intelligence and Legal Liability: Towards an International Approach of Proportional Liability Based on Risk Sharing', (2021) 30 *Information & Communications Technology Law* 169, 175 f.



damages. The liability rules of the current law do not at all provide a convincing solution.<sup>15</sup> There is also no sign of a helpful proposal *de lege ferenda*. In the case of high-frequency trading, this risk has become apparent.<sup>16</sup> As two observers pointedly put it: ‘Who should bear these massive risks of algorithms that control the trading systems, to behave for some time in an uncontrolled and incomprehensible manner and causing a loss of billions?’<sup>17</sup>

## *ii. Big Data*

Incorrect estimates of Big Data analyses cause further liability gaps. Big Data is used to predict how existing societal trends or epidemics can develop and – if necessary – be influenced by vast amounts of data. If the faulty calculation, ie algorithm or underlying data basis, cannot be clearly established, there are difficulties in determining causality and misconduct.<sup>18</sup>

## *iii. Digital Hybrids*

In computational journalism, in other fields of hybrid writing and in several instances of hybrid cooperation, human action and algorithmic calculations are often so intertwined that it becomes virtually impossible to identify which action was responsible for the damage. The question arises of whether liability can be founded on the collective action of the human-machine association itself.<sup>19</sup>

## *iv. Algorithmic Contracts*

An unsatisfactory liability situation arises in the law on contract formation when applied to software agents’ declarations. Once software agents issue legally binding declarations but misrepresent the human as the principal relying on the agent, it is unclear whether the risk is attributed entirely to the principal. Some authors argue that doing so would be an excessive and unjustifiable burden, especially when it comes to distributed action or self-cloning.<sup>20</sup>

<sup>15</sup> So clearly, K Yeung, *Responsibility and AI: A Study of the Implications of Advanced Digital Technologies (Including AI Systems) for the Concept of Responsibility within a Human Rights Framework* Council of Europe study DGI(2019)05, 2019), 62 ff.

<sup>16</sup> eg: M-C Gruber, ‘On Flash Boys and Their Flashbacks: The Attribution of Legal Responsibility in Algorithmic Trading’, in M Jankowska et al. (eds), *AI: Law, Philosophy & Geoinformatics* (Warsaw, Prawa Gospodarczego, 2015) 100.

<sup>17</sup> S Kirn and C-D Müller-Hengstenberg, ‘Intelligente (Software-)Agenten: Von der Automatisierung zur Autonomie? – Verselbstständigung technischer Systeme’, [2014] *Multimedia und Recht* 225, 227 (our translation).

<sup>18</sup> eg: G Kirchner, ‘Big Data Management: Die Haftung des Big Data-Anwenders für Datenfehler’, [2018] *InTeR Zeitschrift zum Innovations- und Technikrecht* 19.

<sup>19</sup> eg: E Dahiyat, ‘Law and Software Agents: Are They “Agents” by the Way?’, (2021) 29 *Artificial Intelligence and Law* 59, 78 ff.

<sup>20</sup> eg: G Sartor, ‘Agents in Cyberlaw’, in G Sartor (eds), *The Law of Electronic Agents: Selected Revised Papers. Proceedings of the Workshop on the Law of Electronic Agents (LEA 2002)* (Bologna, University of Bologna, 2003).

v. *Digital Breach of Contract*

If a contract's performance is delegated to an autonomous software agent and if the agent violates contractual obligations, the prevailing doctrine argues that the rules of vicarious liability for auxiliary persons do not apply. The reason is that an algorithm does not have the necessary legal capacity to act as a vicarious agent. Instead, liability shall only arise when the human principal himself commits a breach of contract. This opens a wide liability gap: once the operator can prove that the software agent has been used correctly without the operator himself having violated a contractual obligation, the operator is not liable.<sup>21</sup> Should the customer then bear the damage caused by the other party's computer?

vi. *Tort and Product Liability*

A similar problem arises in non-contractual liability because, in the case of fault-based liability, it is only the breach of duty prescribed in tort law or product liability law committed by the operator, manufacturer, or programmer that leads to liability. If the humans involved comply with these obligations, then there is no liability.<sup>22</sup> The liability gap will not be closed, even if the courts overstretch duties of care for human actors.<sup>23</sup> The rules of product liability give a certain relief, but they do not close the liability gap. If the decisions of autonomous algorithms cause damage, the injured party will be without protection.

vii. *Liability for Industrial Hazards*

Even legal policy proposals that specify *de lege ferenda* compensation for digital damages with strict industrial hazard liability rules<sup>24</sup> cannot avoid substantial liability gaps. The principles of strict liability can hardly serve as a model since they do not fit the specific risks of digital decisions.

## C. Current Law's Denial of Reality

Liability gaps thus effectively arise when liability law insists on responding to the new digital realities exclusively with traditional concepts that have been developed

<sup>21</sup> See: G Wagner and I Luyken, 'Haftung für Robo Advice', in G Bachmann et al. (eds), *Festschrift für Christine Windbichler* (Berlin, de Gruyter, 2020) 168; MA Chinen, 'The Co-Evolution of Autonomous Machines and Legal Responsibility', (2016) 20 *Vanderbilt Journal of Law & Technology* 338, 363.

<sup>22</sup> This is where authors discover the liability gap for algorithmic acts in product liability law, MA Chinen, *Law and Autonomous Machines* (Cheltenham, Elgar, 2019) 27; G Spindler, 'Zivilrechtliche Fragen beim Einsatz von Robotern', in E Hilgendorf (ed), *Robotik im Kontext von Recht und Moral* (Baden-Baden, Nomos, 2014) 72 ff, 78.

<sup>23</sup> Criticising the trend toward overloading of duties, M-C Gruber, *Bioinformationsrecht: Zur Persönlichkeitsentfaltung des Menschen in technisierter Verfassung* (Tübingen, Mohr Siebeck, 2015) 238 ff.

<sup>24</sup> See prominently: EU Parliament, Resolution 2017, para 6.

for human actors.<sup>25</sup> Adhering to the conventional idea that only human actors dispose of legal subjectivity while seeking to keep pace with the digital developments, legal doctrine is forced to react to the hitherto unknown AI systems with questionable fictions and auxiliary constructions. In the field of contract formation, legal doctrine firmly maintains that only human actors are in the position to make legally binding declarations for them and for others. Therefore, contract law is forced to conceal the independent role of algorithms behind untenable fictions. In the field of contractual and non-contractual liability, damages attributable to a human-computer network must be permanently linked to a negligent damage-causing action of the human actors behind the computer.<sup>26</sup> As a result, it is no longer possible to clearly identify whether all fault-based liability requirements are met. The rules on strict liability lean much too far in one direction but not far enough in another because they treat the digital risk like the mere causal risk of a dangerous object. Finally, there is general perplexity in the legal debate regarding the interconnectivity of algorithmic multi-agent systems.

What is more, legal doctrine attempts to justify its fictions not only by its time-honoured anthropocentric traditions but by a profound humanism that insists that only human beings have the capacity to act. The critique of such an attitude cannot be harsh enough:

A mistaken humanism, blindly complacent and thus deeply inhuman, wants to attribute the behaviour of intelligent machines always and everywhere to human beings, willing to pay the price of any fiction and any doctrinal distortion whatsoever. This is simply ignorant stubbornness, a lack of understanding of technical reality.<sup>27</sup>

Suppose the law continues to react to the use of AI systems – robots, software agents, human-machine-associations, or multi-agent systems – exclusively with traditional concepts tailored for human actors and thus leaves those responsibility gaps unresolved. In that case, it inevitably contributes to damage not being distributed collectively across society, but rather in a merciless *casum sentit dominus* fashion. This is the fundamental reason for massive criticism. Imposing the consequences on the victims who suffered the loss is rightly criticised, both in legal policy terms as well as based on a fundamental sense of fairness. To shield producers and users from responsibility for the damage that unpredictable algorithms cause effectively results in subsidising the most dangerous part of their activities, ie those decisions that escape human control. To qualify them as mere ‘casualties’ that must be borne by their victims, as some suggest,<sup>28</sup> seems almost cynical in

<sup>25</sup> For details, in ch 3, III.B and IV.A, V.A.

<sup>26</sup> eg: C Cauffman, ‘Robo-Liability: The European Union in Search of the Best Way to Deal with Liability for Damage Caused by Artificial Intelligence’, (2018) 25 *Maastricht Journal of European and Comparative Law* 527, 529 f.

<sup>27</sup> P Femia, ‘Soggetti responsabili: Algoritmi e diritto civile’, in P Femia (ed), *Soggetti giuridici digitali: Sullo status privatistico degli agenti software autonomi* (Napoli, Edizioni Scientifiche Italiane, 2019) 9 f (our translation).

<sup>28</sup> eg: M Auer, ‘Rechtsfähige Softwareagenten: Ein erfrischender Anachronismus’, (2019) *Verfassungsblog* 30 September 2019, 5/7 ff.

the light of the new risks that agents' uncontrollable behaviour creates. It is not by chance that the critique of such a cynical attitude comes with particular emphasis from observers of AI-introduction in medical treatment:

The diffusion of responsibility and liability can have problematic consequences: the victim might be left alone, the damages might remain unresolved, and society might feel concerned about a technological development for which accountability for damages and violations of rights remains unclear. Fragile arrangements of trust can break, pre-existing reservations and unease about AI be amplified, and calls for overly restrictive governance result if public attitudes, narratives and perceptions are not taken seriously and channelled into inclusive societal deliberations.<sup>29</sup>

In terms of policy, immunity from liability in these constellations will lead to an oversupply of just those problematic activities.<sup>30</sup> Holding no one liable for unlawful failures of unpredictable algorithms in these hard cases and accepting coincidental losses creates false incentives for operators, producers, and programmers. It will lead to fewer precautions to avoid damage created by the new digital autonomy.<sup>31</sup> Moreover, society's willingness to fully exploit algorithms' promising potential diminishes when the victims have to bear its risks. But also, the mere uncertainty about potential liability has its problems. Above all, however, immunity from liability for digital decisions contradicts a fundamental postulate of justice, demanding a strict connection between decision and responsibility.<sup>32</sup> And the legal principle of equal treatment requires not to privilege users of computers when the same tasks usually delegated to human actors are now delegated to AI systems.

## II. The Overshooting Reaction: Full Legal Subjectivity for E-Persons?

Full legal personhood for autonomous algorithms – this is the much-discussed answer of many lawyers and politicians in the common law world<sup>33</sup> as well as in

<sup>29</sup> M Braun et al., 'Primer on an Ethics of AI-Based Decision Support Systems in the Clinic', (2020) 0 *Journal of medical ethics* 1, 4.

<sup>30</sup> eg: G Wagner, 'Robot Liability', in R Schulze et al. (eds), *Liability for Robotics and in the Internet of Things* (Baden-Baden/Oxford, Nomos/Hart, 2019) 30 f. For the policy arguments of how to deal with robots in an economic perspective, see: A Galasso and H Luo, 'Punishing Robots: Issues in the Economics of Tort Liability and Innovation in Artificial Intelligence', in A Agrawal et al. (eds), *The Economics of Artificial Intelligence: An Agenda* (Chicago, University of Chicago Press, 2019) 495; see generally: S Shavell, *Foundations of Economic Analysis of Law* (Harvard, Harvard University Press, 2004) 208 ff.

<sup>31</sup> eg: H Eidenmüller, 'The Rise of Robots and the Law of Humans', (2017) 27/2017 *Oxford Legal Studies Research Paper* 1, 8.

<sup>32</sup> EU Parliament, Resolution 2017, para 7; reiterated in EU Parliament, Resolution 2020, Proposal for Regulation, Preamble, para 8.

<sup>33</sup> For recent statements, A Lai, 'Artificial Intelligence, LLC: Corporate Personhood as Tort Reform', (2021) 2021 *Michigan State Law Review* Forthcoming, section III.A.; J Turner, *Robot Rules: Regulating*

Continental civil law systems.<sup>34</sup> In January 2017, the European Parliament adopted a resolution based on the Delvaux report that proposed to establish a special legal status for robots and at least grant the most sophisticated autonomous robots the status as ‘electronic persons’ (e-persons) with special rights and obligations, including the redress of all the damage they cause. When robots make autonomous decisions, they should be recognised as ‘electronic persons’, as legal persons in the full sense of the word.<sup>35</sup>

To compensate for the deficiencies mentioned above, several authors have suggested that e-persons should have the ability to make declarations of intent as full legal entities, both in their own name and in the name of others.<sup>36</sup> Moreover, they should be capable of owning property, disposing of money, having bank accounts in their own name and having access to credit. In fact, e-persons are supposed to collect commissions for their transactions and use this self-earned money to pay for damages or infractions.<sup>37</sup> Liability law requires, it is argued, a genuine self-liability of the e-persons: ‘It is possible to hold autonomous agents themselves, and not only their makers, users or owners, responsible for the acts of these agents.’<sup>38</sup> Either the e-persons are allocated a fund for this purpose under property rights, which is alimented by payments from the parties involved (manufacturers, programmers, operators, users), or an insurance policy ought to cover the agent’s own debts.<sup>39</sup>

*Artificial Intelligence* (London, Palgrave Macmillan, 2018) 173 ff; SM Solaiman, ‘Legal Personality of Robots, Corporations, Idols and Chimpanzees: A Quest for Legitimacy’, (2017) 25 *Artificial Intelligence and Law* 155; TN White and SD Baum, ‘Liability for Present and Future Robotics Technology’, in P Lin et al. (eds), *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence* (Oxford, Oxford University Press, 2017); EJ Zimmerman, ‘Machine Minds: Frontiers in Legal Personhood’, (2015) *SSRN Electronic Library* 1.

<sup>34</sup>For Germany: C Kleiner, *Die elektronische Person: Entwurf eines Zurechnungs- und Haftungssubjekts für den Einsatz autonomer Systeme im Rechtsverkehr* (Baden-Baden, Nomos, 2021), 145 ff; D Linardatos, *Autonome und vernetzte Aktanten im Zivilrecht: Grundlinien zivilrechtlicher Zurechnung und Strukturmerkmale einer elektronischen Person* (Tübingen, Mohr Siebeck, 2021) 479 ff; J-P Günther, *Roboter und rechtliche Verantwortung: Eine Untersuchung der Benutzer- und Herstellerhaftung* (Munich, Utz, 2016) 251 ff.

<sup>35</sup>See especially: European Parliament, Resolution 2017, para 18. This prominent European Parliament’s suggestion for recognition of e-persons remained unmentioned in the further European policy debate, already in the responding document outlining the European Strategy on AI by European Commission, Communication ‘Artificial Intelligence for Europe’, COM(2018) 237 final, and were later on not further pursued by the Parliament itself.

<sup>36</sup>eg: J Linarelli, ‘Artificial General Intelligence and Contract’, (2019) 24 *Uniform Law Review* 330, 340 ff; S Wettig and E Zehendner, ‘The Electronic Agent: A Legal Personality under German Law?’, [2003] *Proceedings of the Law and Electronic Agents Workshop* 97, 97 ff.

<sup>37</sup>MU Scherer, ‘Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies’, (2016) 29 *Harvard Journal of Law & Technology* 353, 399.

<sup>38</sup>J Hage, ‘Theoretical Foundations for the Responsibility of Autonomous Agents’, (2017) 25 *Artificial Intelligence and Law* 255, 255; see also: White and Baum, ‘Robotics Technology’ 70 ff.

<sup>39</sup>eg: DC Vladeck, ‘Machines without Principals: Liability Rules and Artificial Intelligence’, (2014) 89 *Washington Law Review* 117, 150; E Hilgendorf ‘Können Roboter schuldhaft handeln? Zur Übertragbarkeit unseres normativen Grundvokabulars auf Maschinen’, in S Beck (ed), *Jenseits von Mensch und Maschine* (Baden-Baden, Nomos, 2012) 127 f.

Beck comes up with a concrete suggestion how to realise full personhood for algorithms:

In practice, this would mean that each such machine would be entered in a public register (similar to the commercial register) and would obtain their legal status at the moment of registration. A change in the owners of the machine's capital stock (most importantly the sale of the machine) should have no impact on the personhood. A certain financial basis would be affixed to autonomous machines, depending on the area of application, hazard, abilities, degree of autonomy, etc. This sum which would have to be raised by the producers and users alike, would be called the 'capital stock' of the robot and collected before the machine was put into public use. The amount of money could also be limited to have an electronic person Ltd. The law should also require a registration number attached to each machine; thus, people interacting with the robot can be informed about the machine's amount of liability, stakeholders, characteristics and other information of the machine.<sup>40</sup>

In private law, they are supposed to become bearers of rights and to assert their own constitutional rights, rights to personal development, non-discrimination, freedom of economic development, and, above all, the right to freedom of expression.<sup>41</sup>

### III. Our Solution: Differential Legal Status Ascriptions for Algorithms

#### A. Algorithms in Social and Economic Contexts

Full legal personhood must be rejected – this is how we argue, together with several authors in common law<sup>42</sup> as well as in civil law<sup>43</sup> and with recent critical EU legal policy perspectives responding to the European Parliament.<sup>44</sup> Demands

<sup>40</sup> S Beck, 'The Problem of Ascribing Legal Responsibility in the Case of Robotics', (2016) 31 *AI & Society* 473, 480. For a thorough discussion of legal structures of e-persons, Linardatos, *Aktanten* 479 ff.

<sup>41</sup> eg: Zimmerman, 'Machine Minds' 34 ff; J Kersten, 'Menschen und Maschinen: Rechtliche Konturen instrumenteller, symbiotischer und autonomer Konstellationen', [2015] *Juristenzeitung* 1, 2 ff, 8.

<sup>42</sup> N Banteka, 'Artificially Intelligent Persons', (2021) 58 *Houston Law Review* 537, 595 f; A Lior, 'AI Entities as AI Agents: Artificial Intelligence Liability and the AI Respondeat Superior Analogy', (2020) 46 *Mitchell Hamline Law Review* 1043, 1067 ff; JJ Bryson et al., 'Of, for, and by the People: The Legal Lacuna of Synthetic Persons', (2017) 25 *Artificial Intelligence Law* 273, 289.

<sup>43</sup> M Ebers, 'Regulating AI and Robots: Ethical and Legal Challenges', in M Ebers and S Navas (eds), *Algorithms and Law* (Cambridge, Cambridge University Press, 2020) 60 ff; R Schaub, 'Interaktion von Mensch und Maschine: Haftungs- und immaterialgüterrechtliche Fragen bei eigenständigen Weiterentwicklungen autonomer Systeme', [2017] *Juristenzeitung* 342, 345 f; N Nevejans, *European Civil Law Rules in Robotics* (Brussels, Study commissioned by the European Parliament's Juri Committee on Legal Affairs, 2016) 14 ff.

<sup>44</sup> Emphatical rejection by the Open Letter to the European Commission, Artificial Intelligence and Robotics, available at [www.robotics-openletter.eu](http://www.robotics-openletter.eu); Expert Group on Liability and New Technologies – New Technologies Formation, Report 'Liability for Artificial Intelligence and Other Emerging