

Law, Governance and Technology Series 10

Ugo Pagallo

The Laws of Robots

Crimes, Contracts, and Torts

 Springer

The Laws of Robots

Law, Governance and Technology Series

VOLUME 10

Series Editors:

POMPEU CASANOVAS, *Institute of Law and Technology, UAB, Spain*

GIOVANNI SARTOR, *University of Bologna (Faculty of Law -CIRSFID)
and European University Institute of Florence, Italy*

Scientific Advisory Board:

GIANMARIA AJANI, *University of Turin, Italy*; KEVIN ASHLEY, *University of Pittsburgh, USA*; KATIE ATKINSON, *University of Liverpool, UK*; TREVOR J.M. BENCH-CAPON, *University of Liverpool, UK*; V. RICHARDS BENJAMINS, *Telefonica, Spain*; GUIDO BOELLA, *Universita' degli Studi di Torino, Italy*; JOOST BREUKER, *Universiteit van Amsterdam, The Netherlands*; DANIELE BOURCIER, *University of Paris 2-CERSA, France*; TOM BRUCE, *Cornell University, USA*; NURIA CASELLAS, *Institute of Law and Technology, UAB, Spain*; CRISTIANO CASTELFRANCHI, *ISTC-CNR, Italy*; JACK G. CONRAD, *Thomson Reuters, USA*; ROSARIA CONTE, *ISTC-CNR, Italy*; FRANCESCO CONTINI, *IRSIG-CNR, Italy*; JESÚS CONTRERAS, *iSOCO, Spain*; JOHN DAVIES, *British Telecommunications plc, UK*; JOHN DOMINGUE, *The Open University, UK*; JAIME DELGADO, *Universitat Politècnica de Catalunya, Spain*; MARCO FABRI, *IRSIG-CNR, Italy*; DIETER FENSEL, *University of Innsbruck, Austria*; ENRICO FRANCESCONI, *ITTIG - CNR, Italy*; FERNANDO GALINDO, *Universidad de Zaragoza, Spain*; ALDO GANGEMI, *ISTC-CNR, Italy*; MICHAEL GENESERETH, *Stanford University, USA*; ASUNCIÓN GÓMEZ-PÉREZ, *Universidad Politécnica de Madrid, Spain*; THOMAS F. GORDON, *Fraunhofer FOKUS, Germany*; GUIDO GOVERNATORI, *NICTA, Australia*; GRAHAM GREENLEAF, *The University of New South Wales, Australia*; MARKO GROBELNIK, *Josef Stefan Institute, Slovenia*; JAMES HENDLER, *Rensselaer Polytechnic Institute, USA*; RINKE HOEKSTRA, *Universiteit van Amsterdam, The Netherlands*; ETHAN KATSH, *University of Massachusetts Amherst, USA*; MARC LAURITSEN, *Capstone Practice Systems, Inc., USA*; RONALD LEENES, *Tilburg Institute for Law, Technology, and Society, Tilburg University, The Netherlands*; PHILIP LIETH, *Queen's University Belfast, UK*; ARNO LODDER, *VU University Amsterdam, The Netherlands*; JOSÉ MANUEL LÓPEZ COBO, *Playence, Austria*; PIERRE MAZZEGA, *LMTG - UMR5563 CNRS/IRD/UPS, France*; MARIE-FRANCINE MOENS, *Katholieke Universiteit Leuven, Belgium*; PABLO NORIEGA, *IIIA-CSIC, Spain*; ANJA OSKAMP, *Open Universiteit, The Netherlands*; SASCHA OSSOWSKI, *Universidad Rey Juan Carlos, Spain*; UGO PAGALLO, *Università degli Studi di Torino, Italy*; MONICA PALMIRANI, *Università di Bologna, Italy*; ABDUL PALIWALA, *University of Warwick, UK*; ENRIC PLAZA, *IIIA-CSIC, Spain*; MARTA POBLET, *Institute of Law and Technology, UAB, Spain*; DANIEL POULIN, *University of Montreal, Canada*; HENRY PRAKKEN, *Universiteit Utrecht and The University of Groningen, The Netherlands*; HAIBIN QI, *Huazhong University of Science and Technology, P.R. China*; DORY REILING, *Amsterdam District Court, The Netherlands*; PIER CARLO ROSSI, *Italy*; EDWINA L. RISSLAND, *University of Massachusetts, Amherst, USA*; COLIN RULE, *University of Massachusetts, USA*; MARCO SCHORLEMMER, *IIIA-CSIC, Spain*; CARLES SIERRA, *IIIA-CSIC, Spain*; MIGEL ANGEL SICILIA, *Universidad de Alcalá, Spain*; RONALD W. STAUDT, *Chicago-Kent College of Law, USA*; RUDI STUDER, *Karlsruhe Institute of Technology, Germany*; DANIELA TISCORNIA, *ITTIG-CNR, Italy*; JOAN-JOSEP VALLBÈ, *Universitat de Barcelon, Spain*; TOM VAN ENGERS, *Universiteit van Amsterdam, The Netherlands*; FABIO VITALI, *Università di Bologna, Italy*; MARY-ANNE WILLIAMS, *The University of Technology, Sydney, Australia*; RADBOUD WINKELS, *University of Amsterdam, The Netherlands*; ADAM WYNER, *University of Liverpool, UK*; HAJIME YOSHINO, *Meiji Gakuin University, Japan*; JOHN ZELEZNIKOW, *University of Victoria, Australia*

For further volumes:

<http://www.springer.com/series/8808>

Ugo Pagallo

The Laws of Robots

Crimes, Contracts, and Torts

 Springer



Ugo Pagallo
University of Torino
Torino Law School
Torino, Italy

ISBN 978-94-007-9804-5 ISBN 978-94-007-6564-1 (eBook)

DOI 10.1007/978-94-007-6564-1

Springer Dordrecht Heidelberg New York London

© Springer Science+Business Media Dordrecht 2013

Softcover reprint of the hardcover 1st edition 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

*To Alexis, Anna Sofia, and the Next
Generation*

Preface

I am inside the orbit of Deimos and completely on my own. Wish me luck!

Curiosity Mars, tweeting on 5 August 2012 at 8:12 p.m., US Pacific Time (two hours and twenty minutes before the robotic rover successfully landed on the red planet).

The year 1961 was notable and moreover, a turning point for one of the most breath-taking fields of today's information revolution, robotics. The amazing pace in the field of robotics and its manifold applications can be traced back to 1961 and the remarkable sequence of events concerning politics, military confrontations, scientific research, culture, society, and the progress of technology. To put things in context, on 12 April 1961, Yuri Gagarin became the first man in space, soon followed by US Navy Commander Alan Shepard on 5 May. In between, about 1300 Cuban exiles armed with US weapons, and sponsored by the CIA, landed at the Bay of Pigs on 17 April, unsuccessfully attempting to overthrow Fidel Castro's regime. Four months later, on 17 August, East Germany (DDR) started to erect the Berlin Wall. A few weeks later, at 11:32 a.m. on 30 October, the USSR detonated the Tsar Bomb, causing the largest man-made explosion in history, namely a 50-megaton hydrogen bomb over the Novaya Zemlya archipelago. Luckily, in this hottest of years during the cold war, technology and science also advanced for more peaceful purposes: Squibb produced the first electric toothbrush, movies were shown for the first time on TWA flights, IBM presented its Selectric typewriter, and Jack Lippes developed the contraceptive intrauterine device. While some glorious movies, such as *West Side Story*, *Breakfast at Tiffany's* and *La Dolce Vita*, were released, a number of unforgettable songs like "Stand by Me," or "Hit the Road Jack," made the charts.

In addition to the publication of such famous books as *Tropic of Cancer* and *The Winter of Our Discontent*, 1961 is also the year when some famous baby boomers were born, such as President Barrack Obama, the jurist Larry Lessig, Princess Diana, George Clooney, Eddie Murphy and, yes, the Fantastic Four: Mister Fantastic, the Invisible Woman, the Human Torch and the monstrous Thing. For that matter, the author of this book was also born in 1961, just in time to enjoy the first disposable diapers, *i.e.*, Pampers.

Besides FM stereos, the new Coca Cola rival of 7 Up, Sprite, and Johnson & Johnson's Tylenol, we should not miss another novelty brought on in 1961. Forty-one years after the word "robot" became popular with Karel Čapek's play *Rossum's Universal Robots* (1920), and almost 20 years after Isaac Asimov coined the term "robotics" in his novel *Runaround* (1942), robots were employed in the industry sector for the first time. Contrarily to Čapek's humanoids and Asimov's artificial agents, these machines were neither robot soldiers nor spacewalkers. Rather, the first industry robot was tested within the automobile sector, drawing on the projects of George Devol and Joseph Engelberger, which culminated in the UNIMATE robot performing spot welding and extracting die-castings in a General Motors factory in New Jersey. Soon after, the idea was not only to manufacture machines (*e.g.*, cars) through further machines (*e.g.*, robots). The plan was to build fully autonomous cars, later dubbed as unmanned ground vehicles, or "UGVs," according to several different projects pursued in the USA, Japan, Germany and Italy.

Yet, it was only 20 years later, in the early 1980s, that the use of robotics within the car industry became critical. Japanese industry first began to implement this technology on a large scale in their factories, acquiring strategic competitiveness by decreasing costs and increasing the quality of their products. As this was the time of my first lengthy stay in Silicon Valley, I vividly recall the sense of shock instilled by the first wave of Japanese automobiles overwhelming Detroit cars on the Californian speedways in summer 1982. Western car producers learned a hard lesson and followed Japanese thinking, installing robots in their factories a few years later. This massive trend went on for two decades: remarkably, in the *Editorial* to the World 2005 Robotics Report of the Economic Commission for Europe and the International Federation of Robotics, Åke Madesäter raised the risk that the robot industry was too focused and dependent on the automotive industry: "The industrial robot industry has become too dominated by car manufacturers and its sub-suppliers. In the period 1997–2003, the automotive industry in Spain received 70 % of all new robot installations. In France, the United Kingdom and Germany the corresponding figure amounted to 68 %, 64 % and 57 %, respectively" (UN 2005: ix).

In the same years as covered by the UN World report, however, things began to rapidly change: the two decade dependence of robotics on the automobile industry dramatically opened up to diversification, a revolution as phrased by scholars. This occurred with water-surface and underwater unmanned vehicles, or “UUVs,” used for remote exploration work and the repairs of pipelines, oil rigs and so on, developing at an amazing pace since the mid-1990s. A decade later, unmanned aerial vehicles (“UAVs”), or systems (“UAS”), upset the military field. As the *U.S. Army Unmanned Aircraft Systems Roadmap 2010–2035* illustrates, their quantitative and qualitative indices are impressive. From 2003 to 2008, UAV flights increased by 2,300 % and the number of UAVs, which was less than 50 before 2001, was over 3,000 in 2006, over 7,000 in 2010, and well over 12,000 at the time of this writing. The impact of UAVs on the laws of war has given rise to UN special rapporteurs and scholars alike proposing stricter regulations for their use. Whereas “the difference between science fiction and science is timing,” in the phrasing of the Colonel Christopher B. Carlile, Director of the UAS Center of Excellence in Fort Rucker, Alabama, it is no surprise then that the Sci-Fi menace of Čapek’s robot soldiers in *R.U.R.* has turned out to be real.

After the UUV and UAV revolutions with their normative challenges, e.g., swarms of tiny drones that plan the missions they are going to execute by themselves, further candidates for the next robotic revolution are a new generation of UGVs, that is, smart cars driving themselves on the highways in fully autonomous, or semi-autonomous, ways. A number of states, organizations and private companies have seriously pursued this project over the past years. Contemplate the Grand Challenge competitions organized by the US Defence Advanced Research Projects Agency (“DARPA”) since the late 1990s. Among the participants of such challenges, suffice it to mention the vehicles sponsored by Carnegie Mellon with General Motors, Stanford with Volkswagen, and Google’s driverless cars. After the Eureka Prometheus Project (1987–1995), the European Commission has similarly promoted the “Intelligent Car Initiative” in 2010, in order to drastically reduce traffic jams and car accidents, while improving energy efficiency and polluting less. Certain terrifying figures can make us fully appreciate that which is at stake with the next UGV revolution: road transport accounts for more than one-quarter of the EU’s total energy consumption, costs of traffic jams amount to approximately 0.5 % of EU GDP, car congestions impact 10 % of the European major road networks, in which there are around 1.3 million mishaps and 41,000 people who die in car accidents every year.

The panoply of robotic applications available suggests further candidates for the next robotic revolution. Reflect on the set of applications for personal

and domestic service: we already have a number of robot toys and robot nannies that are programmed to provide love and take care of children and the elderly. In the academic field, think of a new generation of artificial assistants for university teachers, as a sort of i-Jeeves that could help us schedule conferences, lectures and meetings. By checking the availability and convenience of logistics in accordance with a number of parameters like budget, time efficiency, or weather average conditions, the robot could report its findings back for a decision or, even, determine the steps of the academic tour by directly accepting invitations, booking hotel rooms, flights and so forth. Moreover, we should take into account the class of robotic scientists that may independently discover new knowledge without the need for human intervention, as occurred with “Adam” at Aberystwyth University and the University of Cambridge in 2009, when researchers confirmed that such a robot discovered new evidence about the genomics of the baker’s yeast *Saccharomyces cerevisiae*. Likewise, consider NASA’s mars rover robots and the Science Laboratory flight team: the one-ton, \$ 2.5 billion machine became especially popular on 5 August 2012 when the robot, Curiosity, using a supersonic parachute and a first-of-its-kind “sky crane,” successfully landed on the red planet to discover more about the martial environment and reach places scientists deem as interesting for further study.

Another amazing class of robotic applications concerns hybrids of natural and artificial systems, much as machines that mimic animals and their behaviour. Although nature has required billions of years to refine its own design, so that many of the ideas on animal-like robotic behaviour often outpace the capacities of today’s technology, several interesting projects are on their way: robots that exploit the design choices of multi-objective ant colonies or of brood comb constructions by the stingless bees, up to the development of unmanned micro-drones that fly like an albatross. Whilst hybrids of natural and artificial systems include such applications as nanorobots controlled by muscle cells, or neuroprostheses translating the thought of quadriplegics, the troubles with the computing power of robots have increasingly been addressed by connecting them to a networked repository on the internet, allowing robots to share the information required for object recognition, navigation and task completion in the real world. As part of the *Cognitive Systems and Robotic Initiative* from the European Union seventh framework programme (FP7/2007–2013), this is the aim of the RoboEarth project as a world wide web for robots, that is, a network and database repository where machines can share information and learn from each other about their behaviour and their environment. Avoiding the shortcomings of traditional approaches, such as on-board computers for robots, the goal is to complete a sort of cloud robotics infrastructure with all that is needed to close the loop from robots to RoboEarth to robots.

In addition to further examples, *e.g.*, AI soccer players, what this panoply of robotic applications makes clear is a paramount aspect of the current information revolution, namely the astonishing exponential pace of innovation and technological progress after two decades of a too dependent car industry sector. This acceleration is usually illustrated, or even summed up, with the “Moore’s law,” *i.e.*, the 1965 self-fulfilling prophecy that the computing power of chips would have doubled every 18 months. In addition to the economical, political, and cultural conditions that may favour the use of a certain technology, the almost five decades-long rates of doubling amounts of computation have not only made feasible what simply was impossible few years before, but have opened up new horizons of further technological development. To clarify this point, let me recall a family story that involves one of the most spectacular flops of Apple’s history, that is, the 1992 personal digital assistant Newton. This sort of proto i-Pad with touch-screen and pen-stylus included some applications, as “names,” “dates,” and “notes,” much as simple tools as time zone maps, currency converter, and calculator, that allowed users to gather, manage, and share their information. Contrary to the i-Pad, however, the reason why, at least for my sister and her colleagues, Newton turned out to be a failure mostly depended on the fact that such Apple devices simply arrived 15 years too early and, frankly, were too expensive. Returning to the field of robotics, and by further considering a number of factors such as public research and development (R&D) support, interagency transfers, and growing access to powerful and cheaper software and hardware, we can thus understand a simple truth: whereas each of the initial leaps in the realm of robotics required a 20-year interval, nowadays it seems that almost every year brings about some sort of robotic revolution. From Asimov’s *Runaround* to the current Mars rover machines, a 70-year old story of robotics can be summarized as a classic symphony in four movements.

First, *adagio ma non troppo*: industrial robots were introduced in the manufacturing sector in 1961, that is, almost 20 years after Asimov’s first novel on robotics. Second, *andante con brio*: the use of robots within the car industry became critical in the early 1980s, that is, 20 years after the introduction of the first industrial robot in the automobile field. Third, *ostinato*: in the early 2000s, certain individuals still had the impression that robotics was too dependent on the automobile industry. Fourth, much as at the end of Beethoven’s ninth symphony, *prestissimo, maestoso, molto prestissimo*: both the quantity and quality of robotics applications have somehow spiralled out of control in the past decade, so much so that the exponential curve of advancement in the field of robotics has given rise to certain exaggerations. In light of the new generation of driverless cars, UAVs and UUVs, robotic scientists, hybrids of natural and artificial systems, and so forth,

advocates of the techno-deterministic stance argue that the current information revolution inexorably shapes the destiny of human beings and their societies, so that intelligent machines will succeed humans and we, as a species, could face extinction. Greater than human intelligence, in other words, will emerge through nanobots, artificial intelligence and robotics, as the main contributing factors to this singular event.

However, we do not have to perceive the advancement of robotics as inexorable as the revolutionary movement of the planets, to acknowledge that a number of robotic applications transform and reshape individual and social environments through a new set of constraints and opportunities. The panoply of such robotic applications entails nevertheless a high degree of specialization, suggesting that we should avoid any sort of broad-brush stroke illustration of the topic. Robotics traditionally draws on such disciplines as engineering and cybernetics, artificial intelligence and computer science, physics and electronics, biology and neuroscience, down to the fields of humanities: politics, ethics, economics, law, etc. The extraordinary variety of robotic applications, on one hand, cautions us against generalizations that would inescapably fall short in determining, say, the normative challenges of the field. Whereas, for example, it is likely that drones and other types of autonomous (lethal) weapons mainly affect such fields as international humanitarian and criminal law, other applications, such as da Vinci robot-surgeons, mostly raise matters of contractual obligations and strict liability rules.

On the other hand, the multi-disciplinary nature of robotics suggests that an all-encompassing view of the field far exceeds the capacities of a single scholar. When Massimo Durante and myself were planning a book on legal informatics and the normative challenges of technology in 2011, we finally decided to seek the expertise of several different contributors, who ended up to be more than 20, so as to provide for an adequate portrayal of the subject matter. Although I have been working on different legal topics within robotics in the past years, examining the normative challenges of such fields as the laws of war, contracts, privacy, and tortious liability, is it wise that I now present my own book on the laws of robots? How could a single author deal with such different magnitudes of complexity, as robotics technology and the law?

There are three reasons why I believe the task is possible. First, a relatively strong consensus on how legal systems should govern the design, production and use of robots, through a complex network of concepts, such as agency, accountability, liability, burdens of proofs, responsibility, clauses of immunity, or unjust damages, still exists. In addition, jurists often claim that robotics neither creates nor modifies concepts, principles, and basic rules of the legal field, in accordance with the traditional outlook on law and robotics

that may be coined here as the no new issues-thesis. In light of this popular viewpoint, one of the primary aims of this book is to test the conventional approach to the field, introducing a complex set of concepts, principles, and ways of legal reasoning pertaining to the laws of robots, in connection with Herbert H. Hart's distinction between plain and hard legal cases. As to the former set of legal issues, scholars deal with a complex web of concepts and notions in legal reasoning that yet leave no doubts as to how to apply norms and rules to a certain state of affairs, *e.g.*, cases of responsibility for the robotic behaviour pursuant to the liability model in accomplice cases of criminal law. As to the hard cases of the law, the disagreement among lawyers may regard the meaning of the terms framing the legal question, the ways such terms are related to each other in legal reasoning, or the role of the principles that are at stake in the case. Paradoxically, the fact that a strong consensus still exists in the field of the laws of robots becomes clearer when the behaviour of robots falls within the loopholes of the system, provoking a new generation of hard cases, or necessitating the intervention of lawmakers at both national and international levels. As a result, this book does not intend to offer an all-embracing depiction of today's state of the legal art and, indeed, some relevant fields such as administrative law, or crucial issues, such as data protection, are set aside. Rather, this book focuses on three legal fields, namely criminal law, contracts, and torts, so as to ascertain whether certain robotic applications, such as autonomous lethal weapons or certain types of robo-traders, truly challenge basic pillars of today's legal systems.

Second, by strictly dwelling on the legal side of robotics, instead of the physical, biological, logical, or engineering laws of the discipline, this book aims to prevent some recurring stalemates on definitional issues. Remarkably, scholars still discuss whether the behaviour of robots should properly be considered as "autonomous" and, moreover, what a robot ultimately is, namely a reprogrammable machine operating in a semi- or fully autonomous way, according to the UN World 2005 Robotics Report or, rather, a machine that can make appropriate decisions by perceiving something complex, as advocates of the "sense-think-act" paradigm propose. Such different approaches reverberate on further definitional issues as, for example, the distinction between robots and other artificial agents on the internet. Hence, in order to tackle the complexity of the field, the approach of this book is typically legal, that is pragmatic. What is at stake does not only concern the engineering meaning of such notions, as the autonomy and self-knowledge of robots, in accordance with the ways in which these machines may either interact out there with humans, and other robots, through their on-board computers, or function as robot.txt files on the web, or somewhere in between online and

offline words. Rather, these notions and differences are instrumental in order to understand how these machines can affect current legal systems, much as they do, *pace* the no new issues-thesis, with crimes of negligence intertwined with matters of causation, or new kinds of responsibility for the behaviour of others in the tort law field. On this basis, the aim is to determine whether one right answer is legally at hand, whether legal systems are open to alternative solutions, or political decisions need to be taken. A typical illustration is given by the distinction between autonomous and semi-autonomous weapons in the field of military robotics, and today's debate on whether lethal force should ever be permitted to be fully automated.

Third, I concede that the time in which the intricacies of robotics technology and its impact on legal systems used to fall within the reach of a single scholar is close to an end. To date, jurists have mostly tackled the novelty of the cases induced by robotics technology with the traditional tools of hermeneutics, that is, through an extensive interpretation of the texts, through the use of analogy, the principles of the system, and so forth. In criminal law, for instance, the traditional legal viewpoint conceives robots either as dangerous animals or their use as an ultra-hazardous activity, so that strict liability rules apply to all the circumstances. In the field of contracts, rights and obligations established by artificial agents are generally interpreted through the traditional legal viewpoint of the robots-as-tools approach, so that strict liability rules govern the behaviour of these machines, binding those humans on whose behalf they act, regardless of whether such conduct was planned or envisaged. In tort law, strict liability rules in the field of robotics are most of the time understood by analogy with a party's responsibility for the behaviour of animals, children, or even employees. Yet, the more robotics advances and becomes more sophisticated, the more likely it is that such machines will need a legal regime of their own. Among the solutions proposed in this book, contemplate new forms of accountability for the behaviour of robots in the field of contracts, which mean that, under certain circumstances, only robots would be held liable for damages caused by them. Likewise, consider new forms of responsibility for the behaviour of others, *e.g.*, robots in the field of torts, so that clauses of negligence-based responsibility could replace some of today's strict liability rules in cases where third parties are the least-cost avoider of the risk. At the end of the day, the aim of this work is not only to pinpoint those principles, norms, and concepts of today's legal systems which are under stress: the purpose is also to take sides before the hard cases of the law as induced by a novel generation of robotic applications. All in all, I think that some types of robots should not be considered as simple tools of human interaction but, rather, proper agents in the legal field.

However, the more robots require a legislation of their own, the more a new team of experts in robotic crimes, pacts and contracts, administrative procedures, copyright and privacy issues, laws of war, torts, and so on, will supersede the efforts of the single scholar. The process of specialization that has occurred within such fields as IT law, or legal informatics, throughout the 2000s, will likely resurface in the field of legal robotics in a few years. Retrospectively, this work is placed at a turning point of the contemporary legal systems, that is, so to speak, between a “not yet” and an “any longer.” Not yet, because a number of challenges brought on by robotic technology and its manifold applications are still open to alternative solutions in the legal domain; any longer, because traditional legal outlooks increasingly fall short in coping with the novelty of such challenges. Let us grasp why we are facing such an in-between state of art in the laws of robots, throughout the chapters of this volume.

Torino, Italy

Ugo Pagallo

Acknowledgments

This book is the final step of a 4-year project (2009–2013). The preliminary stages were the papers and articles that I have been discussing and publishing over the past years. First, thanks are given to the referees and editors of the journals and books where I published my previous robotic works, a detailed list of which is given below in the references. In particular, let me thank Mariarosaria Taddeo, who edited the special issue of *Knowledge, Technology & Policy* (2010: 23) on “Trust in Technology,” in which I presented “Robotrust and Legal Responsibility” (Pagallo 2010a); Terry Bynum and Simon Rogerson, the founders and souls of the Ethicomp meetings, where I delivered “The Human Master with a Modern Slave?” (Pagallo 2010b), and “The Adventures of Picciotto Roboto” (Pagallo 2011a); Greg Michaelson and Ruth Aylett, editors of the special issue of *AI & Society* (2011: 26(4)) on the “Social Impact of AI: Killer Robots or Friendly Fridges,” with my “Killers, Fridges, and Slaves” (Pagallo 2011b); John Sullins, who edited the special issue of *Philosophy & Technology* (2011: 24(3)) on “Open Questions in Roboethics,” with my “Robots of Just War” (Pagallo 2011c); Herman Tavani, editor with Dieter Arnold of the special issue of *Information* (2011: 2(2)) on “Trust and Privacy in Our Networked World,” where “Designing Data Protection Safeguards Ethically” was published (Pagallo 2011d); Brendan Gogarty, who invited me to deliver an expert commentary for the special edition of the *Journal of Law, Information and Science* (2011) on “Laws Unmanned,” that is, my paper on “Guns, Ships, and Chauffeurs” (Pagallo 2011e); and, last but not least, Mireille Hildebrandt, who edited with Jeanne Gaakeer the Springer volume on “Human Law and Computer Law,” with my essay on “What Robots Want” (Pagallo 2013).

All this previous work represents the starting blocks of this volume, together with both the papers for the AICOL series, coedited with Gianmaria

Ajani, Pompeu Casanovas, Monica Palmirani, and Giovanni Sartor (Pagallo 2010c, 2012a); and the entry “Robotica” for the UTET volume on legal informatics, coedited with Massimo Durante (Pagallo 2012b). During the Fall semester of 2011, spent at the University of Uppsala, a first draft of this book was completed thanks to the formal revision and substantial remarks of Patricia Mindus and Laura Carlson. The manuscript was revised a second time during the Spring semester of 2012, spent at my own university in Turin, where I delivered my course on legal informatics and robotics. A number of colleagues and friends should be thanked for their support, much as the students of my course for their questions and theoretical curiosity. Together with Gianmaria Ajani and Massimo Durante, let me especially thank Raffaele Caterina and Michele Graziadei. By the end of April 2012, I then followed the advice of Greg Chaitin: after pinning the book down as slowly as possible, I let it rest for a while. Three months later, in August 2012, a third revision was undertaken and the preface completed in Cupertino, CA. During the wonderful weeks spent in my favourite villa, I enjoyed the further insights of an eminent expert in machine learning and AI, namely my sister Giulia, and of a distinguished mathematician, my brother-in-law Victor Pereyra.

Diachronically, the book was also improved or, at least, some of its limits and vagueness superseded thanks to many conversations with Luciano Floridi, with whom I had the honour to be member of the group of experts on “the onlife initiative,” set up by the European Commission as part of the Digital Futures project in 2012. The final revision of the book was completed in January 2013, paying attention to the remarks of the reviewers, much as the suggestions of further colleagues and friends. Among them, let me mention Chuck Abernathy from Georgetown University for his common law wisdom. As to the practical side of this volume, thanks are given to the editors of the Springer series on “Law, Governance and Technology,” that is Pompeu Casanovas and Giovanni Sartor, together with Neil Olivier, Senior Publishing Editor of Springer, and his assistant Diana Nijenhuijzen. From the initial project of the book in August 2011 to the green light of the Springer team by the end of winter 2012–2013, all of them helped me to make that August ’11 project real.

Despite this manner of production and the number of inputs by reviewers, colleagues, and friends, I am conscious that the book may still have ambiguities, imprecisions, or simply mistakes. This possibility reminds me of the introductory scene of Čapek’s *Rossum’s Universal Robots*, where Domin, the General Manager of R.U.R., explains to Helena that they were building robots by the thousands, able to speak, write and do arithmetic, free from errors and with a formidable memory. This original idea has significantly fed popular beliefs ever since, down to the point that a pop song

recently reminds us that “I am not a robot.” Although one of the most critical issues of robotics concerns the degree of their error-proneness, let alone whether these machines could have some types of emotions, like falling in love, this naïf version of the field functions as a proemial warning. The continuous process of reviewing and the suggestions of colleagues and friends helped me to improve the previous versions of this book and, yet, some imperfections may still remain. Updating Augustin of Hippo’s proverb, to err is human, to persist is of the bad robotic designer.

Contents

1	Introduction	1
2	On Law, Philosophy and Technology	19
2.1	The Philosophy of Law and Robots	21
2.1.1	The Law in Literature	22
2.1.2	Sources, Concepts, and Legal Reasoning	25
2.1.3	The Levels of Abstraction	28
2.2	The Principle of Responsibility	29
2.2.1	Immunity	31
2.2.2	Strict Liability	33
2.2.3	Personal Fault	34
2.2.4	Responsibility for a Robot	35
2.3	Agency and Accountability of Artificial Agents	37
2.3.1	A Moral Threshold	38
2.3.2	Agents Before the Law	40
2.4	Who Pays?	43
3	Crimes	45
3.1	Sci-Fi Scenarios	49
3.2	The States of Mind and Criminal Acts	52
3.3	Robots and Just Wars	55
3.3.1	What Robots Might Change	57
3.3.2	Just Causes of War	58
3.3.3	Conditions of Just Wars	60
3.3.4	Proportionality	62
3.4	The Phenomenology of <i>Picciotto Roboto</i>	65
3.4.1	<i>Picciotto</i> by Design	66
3.4.2	Crimes of Intent	69
3.4.3	Crimes of Negligence	71
3.5	A Failure of Causation?	73

- 4 Contracts**..... 79
 - 4.1 Pacts, Clauses and Risk..... 83
 - 4.2 The Artificial Doctor 88
 - 4.2.1 Parties, Counterparties and Third Parties 89
 - 4.2.2 Producers, Users and Patients 91
 - 4.3 Robo-Traders 95
 - 4.3.1 Artificial Greediness 96
 - 4.3.2 The Robot and the Principal 97
 - 4.3.3 A New Agent in Town..... 101
 - 4.4 Modern Robots, Ancient Slaves 102
 - 4.4.1 The Digital Peculium 103
 - 4.5 The UV Revolution 106
 - 4.5.1 AI Chauffeurs and Intelligent Car Sharing..... 108
 - 4.5.2 Unjust Damages 111
- 5 Torts** 115
 - 5.1 Bad Intentions 119
 - 5.2 Children, Pets and Negligence 121
 - 5.2.1 American Parents 124
 - 5.2.2 Italian Parents 126
 - 5.3 AI Employees and Strict Liability Rules 130
 - 5.3.1 The Digital Peculium Revisited 132
 - 5.4 Burdens of Proof 135
 - 5.4.1 The Precautionary Principle 138
 - 5.4.2 Robotic Openness 143
- 6 Law as Meta-technology** 147
 - 6.1 Robots as Legal Persons 152
 - 6.1.1 The Front of Robotic Liberation 155
 - 6.1.2 The Pragmatic Stance 163
 - 6.2 Robots as Strict Agents 166
 - 6.3 Sources of Good and Evil 170
 - 6.4 Levels of Complexity 174
 - 6.4.1 Technologies of Social Control 177
 - 6.4.2 The Political Requirement 179
- Conclusions**..... 183
- References**..... 193

List of Figures

Fig. 1.1	The magnitudes of complexity of robotics technology	7
Fig. 1.2	A philosophy of law for lawyers and a work in positive law for philosophers	10
Fig. 1.3	Three legal fields for responsible robots	14
Fig. 2.1	Levels of abstraction	22
Fig. 2.2	A first model for the philosophy of law and robots.....	25
Fig. 2.3	A second model for the philosophy of law and robots.....	27
Fig. 2.4	A new interface for the philosophy of law and robots	28
Fig. 2.5	Three conditions of responsibility for the construction and use of robots	31
Fig. 2.6	From responsibility to legal agency and return.....	41
Fig. 3.1	The Phenomenology of Picciotto Roboto, step 1	67
Fig. 3.2	Phenomenology of Picciotto Roboto, step 2.....	70
Fig. 3.3	Phenomenology of Picciotto Roboto, step 3.....	72
Fig. 4.1	Contractual obligations and robotics complexity.....	81
Fig. 5.1	A common law approach to negligence in the law of Torts.....	125
Fig. 5.2	A civil law approach to the law of Torts	127
Fig. 5.3	Strict liability for robots in the law of Torts.....	130
Fig. 5.4	Reversing the burden of proof with the precautionary principle	139
Fig. 6.1	Law and the challenges of technology	149
Fig. 6.2	Levels of legal complexity in the governance of robotics....	175
Fig. 6.3	Four robotic challenges to law as meta-technology	178
Fig. A.1	Three roads to design	184
Fig. A.2	A teleological approach to design.....	186

List of Tables

Table 1.1	The behaviour of robots and nine ideal-typical conditions of legal responsibility	13
Table 4.1	What the approach to robots-as-tools lacks	100
Table 6.1	Robots' behaviour and the "Factual Limits" of legal science	165
Table 6.2	A threshold of robots' responsibility in the civil law field	169
Table 6.3	The challenges of today's laws of robots as a source of damage	173

Law, Governance and

Ugo Pagallo

The Laws of Robots

Crimes, Contracts, and

Biblioteka Główna
Akademii Sztuki Wojennej

26677/III (CB)



03-026677-000-0

This book explores how the design, construction, and use of robotics technology may affect today's legal systems and, more particularly, matters of responsibility and agency in criminal law, contractual obligations, and torts. By distinguishing between the behaviour of robots as tools of human interaction, and robots as proper agents in the legal arena, jurists will have to address a new generation of "hard cases." General disagreement may concern immunity in criminal law (*e.g.*, the employment of robot soldiers in battle), personal accountability for certain robots in contracts (*e.g.*, robo-traders), much as clauses of strict liability and negligence-based responsibility in extra-contractual obligations (*e.g.*, service robots in tort law). Since robots are here to stay, the aim of the law should be to wisely govern our mutual relationships.

Law

ISBN 978-94-007-9804-5



9 "789400"798045"

► springer.com

