The Cambridge Handbook of
ARTIFICIAL INTELLIGENCE

Global Perspectives on Law and Ethics

EDITED BY

Larry A. DiMatteo, Cristina Poncibò
and Michel Cannarsa

CAMBRIDGE

THE CAMBRIDGE HANDBOOK OF ARTIFICIAL INTELLIGENCE

The technology and application of artificial intelligence (AI) throughout society continues to grow at unprecedented rates, which raises numerous legal questions that to date have been largely unexamined. Although AI now plays a role in almost all areas of society, the need for a better understanding of its impact, from legal and ethical perspectives, is pressing, and regulatory proposals are urgently needed. This book responds to these needs, identifying the issues raised by AI and providing practical recommendations for regulatory, technical, and theoretical frameworks aimed at making AI compatible with existing legal rules, principles, and democratic values. An international roster of authors including professors of specialized areas of law, technologists, and practitioners bring their expertise to the interdisciplinary nature of AI.

LARRY A. DIMATTEO is Huber Hurst Professor of Contract Law at the Warrington College of Business and Levin College of Law, University of Florida. He was the University of Florida's 2012 Teacher-Scholar of the Year and is the former Editor-in-Chief of the *American Business Law Journal*. He is the author, coauthor, or coeditor of more than 150 publications, including 15 books. His books include *The Cambridge Handbook of Judicial Control of Arbitral Awards* (edited; Cambridge University Press, 2020); *The Cambridge Handbook of Smart Contracts, Blockchain Technology and Digital Platforms* (edited; Cambridge University Press, 2019); *Comparative Contract Law: British and American Perspectives* (edited; Oxford University Press, 2nd ed., 2021); and *International Sales Law: Principles, Contracts and Practice* (edited; Beck, Hart, & Nomos, 2016).

CRISTINA PONCIBÒ is Professor of Comparative Private Law at the Law Department of the University of Turin, Collegio Carlo Alberto Affiliate and a faculty member at Georgetown Law Center for Transnational Legal Studies, London. She is also a fellow of the Transatlantic Technology Law Forum (Stanford Law School and Vienna School of Law). Her most recent books include *Contracting and Contract Law in the Age of Artificial Intelligence* (edited; Hart, 2022) and *The Cambridge Handbook of Smart Contracts, Blockchain Technology and Digital Platforms* (edited; Cambridge University Press, 2019). She is the scientific director of the Master's in International Trade Law at the University of Turin, ITC-ILO, in cooperation with Unicitral and Unidroit. In her career, she has been a Marie Curie Intra-European fellow (Université Panthéon-Assas) and a Max Weber fellow (European University Institute).

MICHEL CANNARSA is Dean of Law at Lyon Catholic University, France. His areas of research are international and European law, commercial law, comparative law, consumer law, law of obligations, and legal translation. His recent works have focused on the interaction between law and technology, contract, and products liability law, including *The Cambridge Handbook of Smart Contracts, Blockchain Technology and Digital Platforms* (Cambridge University Press, 2019); "Interpretation of Contracts and Smart Contracts," *European Review Private Law* (2018); "Remedies and Damages," in *Chinese Contract Law: Civil and Common Law Perspectives* (DiMatteo and Lei, eds., Cambridge University Press, 2017); and *La responsabilité du fait des produits défectueux: étude comparative* (Giuffrè, 2005). He is a fellow of the European Law Institute.

The Cambridge Handbook of Artificial Intelligence

GLOBAL PERSPECTIVES ON LAW AND ETHICS

Edited by

LARRY A. DIMATTEO

University of Florida

CRISTINA PONCIBÒ

University of Turin

MICHEL CANNARSA

Lyon Catholic University

Contents

Figures

Contributors

Pınar Çağlayan Aksoy is an associate professor of Civil Law at Bilkent University Faculty of Law, Ankara. She received her LLB degree from Bilkent University Faculty of Law and her LLM and PhD from Ankara University Faculty of Law. She is a member of Blockchain Turkey Platform, Istanbul Blockchain Women Society, and the University of Zurich Blockchain Center. As a member of the European Law Institute, she acts as a member of the Consultative Committee for the project titled "Blockchain Technology and Smart Contracts." Her current research deals with contract law, tort law as well as the legal implications of the newly emerging technologies, especially artificial intelligence and distributed ledger technology. Çağlayan Aksoy has authored and edited books and published articles on international sales law, comparative contract law as well as tort law.

Sam Boro practices in Washington, DC, where he counsels clients including banks, fintech companies, financial services firms, and credit card companies on payments processing, consumer protection, business operations, government investigations, and regulatory compliance. His primary focus is technology transactions and privacy law in such industries as artificial intelligence, machine learning, payments, blockchain, digital assets, and fintech.

Michel Cannarsa is Dean of Law at Lyon Catholic University. His areas of research are international and European law, commercial law, comparative law, consumer law, law of obligations, and legal translation. His recent works have focused on the interaction between law and technology, contract, and products liability law, including *The Cambridge Handbook of Smart Contracts, Blockchain Technology and Digital Platforms* (Cambridge University Press, 2019); "Interpretation of Contracts and Smart Contracts: Smart Interpretation or Interpretation of Smart Contracts?," *European Review Private Law* (2018); "Remedies and Damages," in *Chinese Contract Law, Civil and Common Law Perspectives* (DiMatteo and Lei, eds.; Cambridge University Press, 2017); and *La responsabilité du fait des produits défectueux: étude comparative* (Giuffrè, 2005). He is a fellow of the European Law Institute.

Giuseppe Colangelo is Jean Monnet Professor of European Innovation Policy and Associate Professor of Law and Economics at the University of Basilicata, and a Transatlantic Technology Law Forum fellow at Stanford University Law School. He is Adjunct Professor of Markets, Regulations, and Law and of Competition and Markets of Innovation at Luiss and Scientific Coordinator of the Research Network for Digital Ecosystem, Economic Policy, and Innovation (Deep-In). His research areas include antitrust, intellectual property, market regulation, digital platforms, and law and economics.

Joshua P. Davis is a professor at the University of San Francisco, School of Law and the director of the Center for Law and Ethics. He has published dozens of articles and book chapters on various topics including artificial intelligence, ethics, and law; jurisprudence; jurisprudence and legal ethics; class action doctrine; private antitrust enforcement; civil procedure; and free speech doctrine. He has testified before Congress on federal civil procedure. He also served as the Reporter for the California Supreme Court Advisory Task Force on Multijurisdictional Practice

and the Committee on Multijurisdictional Practice, drafting rules on multijurisdictional practice that were codified at California Supreme Court Rules 964–967. His publications have been cited by federal trial and appellate courts.

Larry A. DiMatteo is the Huber Hurst Professor of Contract Law at the Warrington College of Business and Levin College of Law at the University of Florida. He was the University of Florida's 2012 Teacher-Scholar of the Year; former Editor-in-Chief of the *American Business Law Journal*; and a 2012 Fulbright Scholar (University of Sofia). He is the author, coauthor, or coeditor of more than 150 publications, including 15 books. His books include *The Cambridge Handbook of Judicial Control of Arbitral Awards* (coedited; Cambridge University Press, 2020); *The Cambridge Handbook of Smart Contracts, Blockchain Technology and Digital Platforms* (coedited; Cambridge University Press, 2020); *Chinese Contract Law: Civil and Common Law Perspectives* (coedited; Cambridge University Press, 2017); *Comparative Contract Law: British and American Perspectives* (coedited with Hogg; Oxford University Press, 2016); *International Sales Law: Principles, Contracts and Practice* (coedited with Janssen, Schulze, and Magnus; Beck, Hart, & Nomos, 2nd ed., 2021); and *Global Challenge of International Sales Law* (edited; Cambridge University Press, 2014).

Luigi Distefano is an associate with Finnegan Europe LLP. With more than ten years of intellectual property experience, Distefano has gained a wide spectrum of technical expertise, particularly in the fields of telecommunications, AI, cryptography, optical devices, computer graphics, cybersecurity, and quantum computing. He is a UK and European patent attorney whose practice focuses on patent prosecution, portfolio development and management, drafting, client counseling, and oppositions and appeals before the European Patent Office. He specializes in the fields of electronic and mechanical engineering and in prosecuting computer-implemented inventions in Europe.

Mateja Durovic is a reader in Contract and Commercial Law and Deputy Director of the Centre for Technology, Ethics, Law and Society at King's College London. Durovic holds PhD and LLM degrees from the European University Institute (EUI), an LLM degree from the University of Cambridge, and an LLB degree from the University of Belgrade. He was a postdoctoral research associate at the EUI, as well as being visiting scholar at Stanford Law School and the Max Planck Institute of Private International and Comparative Law. Durovic has worked for the Legal Service of the European Commission, as well as being a consultant for the European Commission, BEUC, and the United Nations. His work has been published in leading law journals and prominent publishers, such as Oxford University Press and Hart Publishing. He is a member of the European Law Institute, Society of Legal Scholars, and Society for European Contract Law.

Martin Ebers is an associate professor of IT Law at the University of Tartu and permanent research fellow at the Humboldt University of Berlin. He is the cofounder and president of the Robotics & AI Law Society. In addition to research and teaching, he has been active in the field of legal consulting for many years. His main areas of expertise and research are IT law, liability and insurance law, and European and comparative law. In 2020, he published the books *Algorithms and Law* (Cambridge University Press) and *Rechtshandbuch Künstliche Intelligenz und Robotik* (C. H. Beck).

Mark Fenwick is Professor of International Business Law at the Faculty of Law, Kyushu University. His primary research interests are white-collar and corporate crime, as well as

technology and law. Recent publications include *Legal Tech, Smart Contracts & Blockchain* (coedited with Corrales and Haapio; Springer, 2019), *Robotics: AI and the Future of Law* (coedited; Springer, 2018); and *International Business Law: Emerging Fields of Regulation* (coauthored with Corrales and Forgo; Hart, 2018). He has a master's and PhD from the Faculty of Law, University of Cambridge, and has been a visiting professor at the University of Cambridge, Chulalongkorn University, Duke University, the University of Hong Kong, Shanghai University of Finance & Economics, National University of Singapore, Tilburg University, and Vietnam National University. He has also conducted research for the EU, the OECD, and the World Bank.

Nicholas Fox is a partner at Finnegan Europe LLP. He practices intellectual property law with a focus on electronics, telecoms, and software patent litigation. With degrees in mathematics and computer science, Fox has technical experience in the areas of electronics and software. The subject matter of his work ranges from image processing and voice recognition to telecoms and e-commerce. He is a qualified solicitor and qualified as a European and Chartered British patent attorney, with full rights to appear in proceedings in the English High Courts. He has also been admitted as an attorney in New York.

Florence G'sell is a professor of Private and Comparative Law at the University of Lorraine and a lecturer at Sciences Po. She holds the Digital, Governance, and Sovereignty Chair at the Sciences Po School of Public Affairs. She has researched in the areas of causation in tort law and the evolution of the judicial system and legal professions. In recent years, her publications have mainly focused on technology. She recently edited the book *Law and Big Data* (Dalloz, 2020) and published *Justice Numérique* (Dalloz, 2021).

Robert A. Heverly is an associate professor of Law at Albany Law School and an affiliated fellow with the Information Society Project at Yale Law School. He holds a JD from Albany Law School and an LLM from Yale Law School. He has served as the director of the LLM in Information, Technology, and Intellectual Property at the University of East Anglia. Heverly's research areas are in technology, law, and society, including drones, robots, AI, and human augmentation. He has held the position of Chair of the American Association of Law Schools' Internet and Computer Law Section and was the Reporter for the Uniform Law Commission's "Uniform Tort Law Relating to Drones Act."

Charlyn L. Ho is a partner at Perkins Coie, Washington, DC. She provides counsel to clients on legal issues related to technology and privacy, including those affecting e-commerce sites, mobile devices and applications, AI/machine learning, virtual reality, mixed reality and augmented reality platforms, cloud services, enterprise software, cryptocurrency platforms, and Internet of Things devices. Ho provides strategic advice and counseling to all types of technology companies throughout their life cycle, from start-ups to established enterprises. She previously served as an active-duty Navy Supply Corps officer and was responsible for negotiating contracts for nuclear-powered aircraft carrier refueling and overhaul projects, such as replacing spent nuclear fuel in nuclear-powered ships.

Geraint Howells is Executive Dean and Established Professor at the University of Galway, and Visiting Professor at the University of Manchester. He was previously Professor of Commercial Law and Associate Dean for Internationalisation for Humanities at the University of Manchester and Chair Professor of Commercial Law and Dean of the Law School at City University of Hong Kong, as well as having held chairs at Sheffield, Lancaster, and Manchester, and being

head of law schools at Lancaster and Manchester. Howells is a barrister at Gough Square Chambers, London and a former president of the International Association of Consumer Law.

Coimbra Jackson practices in Washington, DC, where she is an associate in the Technology Transactions & Privacy Law practice. Jackson advises clients on issues relating to technology-related transactions. Her clients include sophisticated and emerging companies offering cloud-based services, Internet of Things products and devices, consumer products, and telecommunications. Jackson has counseled on legal issues arising from artificial intelligence and machine-learning products, including issues related to privacy and machine-learning product liability. She is currently engaged in work focused on combating bias in artificial intelligence and machine learning in consumer-facing services.

André Janssen is Chair Professor at the University of Radboud (Nijmegen) and holds the Francqui-chair at the Catholic University Leuven. He previously held positions at the Universities of Münster, Turin and City University of Hong Kong. He has published more than 170 books and articles in the field of private, European and comparative law and artificial intelligence and law. His latest books are *International Sales Law: Contract, Principles & Practice* (coedited with DiMatteo, Magnus, and Schulze; Beck, Hart, & Nomos, 2nd ed., 2021) and the *Cambridge Handbook of Lawyering in the Digital Age* (coedited with L. DiMatteo, P. Ortolani, F. de Elizalde, M. Cannarsa and M. Durovic, CUP, 2021). He is the chief editor of the *European Review of Private Law* and a member of the editorial board of the *International Arbitration Law Review*.

Jonas Knetsch is a professor of Civil and Comparative Law at the Sorbonne Law School at the University Paris 1 Panthéon-Sorbonne. Before his PhD in Comparative Tort Law, he graduated from the University Panthéon-Assas (Paris 2) and from the University of Cologne. He previously held professorships at the University of Reunion Island and at the Jean Monnet University of Saint-Étienne. Knetsch has published extensively in tort and insurance law as well as in private international and comparative law. He is a fellow of the European Centre of Tort and Insurance Law and has an associate membership of the International Academy of Comparative Law.

Pascal D. König is a research associate at the chair of Political Science with a focus on Policy Analysis and Political Economy at TU Kaiserslautern, Germany. His research mainly deals with policies regarding digital technologies, political communication, and party competition. Recent work has appeared in Comparative Political Studies, Big Data & Society, and Regulation & Governance.

Tobias D. Krafft is a PhD candidate at the chair "Algorithm Accountability" of Prof. Katharina A. Zweig at the TU Kaiserslautern. He is an expert in black box analysis and regulatory approaches for algorithmic decision systems. As holder of the Weizenbaumpreis 2017 of the Forum Informatiker für Frieden und gesellschaftliche Verantwortung, his research interests range from the (pure) analysis of algorithmic decision systems to discourse on their use in the social context.

Migle Laukyte is a tenure track professor in Cyberlaw and Cyber Rights at the University Pompeu Fabra in Barcelona. Previously she was a visiting professor at the Human Rights Institute "Bartolomé de las Casas" at the Universidad Carlos III de Madrid and CONEX-Marie Curie fellow at ALLIES (Artificially Intelligent Entities: Their Legal Status in the Future), specifically drafting a project for a model of legal personhood for AI. She earned her PhD at the Bologna University School of Law and was a Max Weber postdoctoral fellow at the

European University Institute. Her research interests are legal, ethical, and philosophical questions related to AI, robotics, and other disruptive technologies. Her recent publications include "The Intelligent Machine: A New Metaphor through which to Understand Both Corporations and AI," *AI and Society* (2020) and "Robots: Regulations, Rights and Remedies," in *Legal Regulations, Implications and Issues Surrounding Digital Data* (Jackson and Shelly, eds.; IGI Global, 2020).

Christian E. Mammen is an intellectual property (IP) litigation partner with Womble Bond Dickinson in Palo Alto, CA. He holds a JD from Cornell Law School and a DPhil in legal philosophy from Oxford University. Mammen has practiced in San Francisco and Silicon Valley for over twenty years and has held visiting faculty positions at Oxford University, UC Berkeley Law School, and UC Hastings College of the Law. He sits on the Berkeley Center for Law & Technology Advisory Board, the Silicon Valley Advanced Patent Law Institute Steering Committee, Law 360's 2020–2022 Intellectual Property Editorial Advisory Board, and the Intellectual Property Owners Association's AI and Emerging Technologies Committee. He is also a cofounder of the Oxford Entrepreneurs Network.

Marc Martin is a partner at Perkins Coie, Washington, DC. He is the chair of the Communications Industry Group and specializes in the areas of regulatory, transactional, and strategic advice to technology, media, and telecom companies, entrepreneurs, investors, and enterprise customers. Martin structures and negotiates agreements in the areas of technology and content licensing, supply chain procurement, mobile apps, and distribution platforms.

John O. McGinnis is the George Dix Professor in Constitutional Law at the Northwestern Pritzker School of Law. He is a graduate of Harvard College and Harvard Law School where he was an editor of the *Harvard Law Review*. He also has an MA degree from Balliol College, Oxford, in Philosophy and Theology. McGinnis clerked on the US Court of Appeals for the District of Columbia. From 1987 to 1991, he was the deputy assistant attorney general in the Office of Legal Counsel at the Department of Justice. He is the author of *Accelerating Democracy: Transforming Government through Technology* (Princeton University Press, 2013) and *Originalism and the Good Constitution* (with Rappaport; Harvard University Press, 2013). He is a past winner of the Paul Bator award given by the Federalist Society to an outstanding academic under forty.

Francesco Mezzanotte is an associate professor at Roma Tre. He is the author of more than fifty publications, including "Risk Allocation and Liability Regimes in the IoT," in *Digital Revolution: New Challenges for Law* (Schulze and De Franceschi, eds.; Beck/Nomos, 2019) and "Access to Data: The Role of Consent and the Licensing Scheme," in *Trading Data in the Digital Economy: Legal Concepts and Tools* (Lohsse, Schulze, and Staudenmayer, eds.; Hart/Nomos, 2017).

Eliza Mik holds a PhD in Contract Law from the University of Sydney. She has taught courses in contract law and in the law of e-commerce at the Singapore Management University and the University of Melbourne, as well as courses in fintech and blockchain at Bocconi University in Milan. Her research focuses on distributed ledger technologies and smart contracts, particularly the challenges of expressing agreements in code as well as domain-specific programming languages. She is also involved in multiple projects relating to the legal implications of automation, the deployment of "intelligent agents" in transactional environments, as well as the burgeoning area of LegalTech. Mik holds multiple academic affiliations, including with the

Tilburg Institute for Law, Society, and Technology, the Center for AI and Data Governance in Singapore, and the Nanjing University of Science and Technology. Before joining academia, she worked in-house at several software companies, internet start-ups, and telecommunication providers in Australia, Poland, Malaysia, and in the United Arab Emirates where she advised on technology procurement, payment systems, and software licensing.

Marjolaine Monot-Fouletier is Professor of Public Law, Director of the Law Clinics, and Head of the Legal, Political, and Social Sciences Research Center at the Faculty of Law of Lyon Catholic University. She holds a PhD in Law (1999) and a habilitation to conduct research from the University Paris-Cité (2019). Her research focus is in the areas of administrative law, administrative property law, and new technologies law. Monot-Fouletier has published over twenty scientific articles, as well as a manual on administrative property law. She is a contributor to the *Encyclopedia of Jurisclasseur* and is a member of the European Law Institute.

Yelena Morozova is an associate with Finnegan Europe LLP. She is an experienced patent attorney with a background in computer science and technical expertise especially in the areas of information technology and software, and telecommunications. Her experience includes patent drafting and prosecution of patent and design applications internationally, developing and implementing intellectual property strategies tailored to business plans, worldwide patent portfolio management, freedom-to-operate searches and infringement, and validity advice. Morozova is a qualified patent attorney in the United States and Europe.

Florian Möslein is Director of the Institute for Law and Regulation of Digitalisation and Professor of Law at the Philipps-University Marburg. Möslein holds academic degrees from the University of Munich, University of Paris-Assas (*licence en droit*), and University of London (LLM in International Business Law). His current research focus is on regulatory theory, corporate sustainability, and the legal challenges of the digital age. He previously held academic positions at the Universities of Berlin, St. Gallen, and Bremen, and visiting fellowships at the European University Institute, NYU, Stanford, Berkeley, University of Sydney, CEU San Pablo, and Aarhus. Möslein has over eighty publications including three monographs and seven edited books.

Christy Ng is a legal technologist who has advised on a range of documentation and innovation initiatives within the derivatives industry and fintech space. She is a consultant with D2 Legal Technology LLP and its Hong Kong subsidiary where she works on projects relating to legal data for machine learning and digitalization of financial services. She has worked with banks, investment firms, asset managers, and law firms in both an external and in-house capacity. Through her experience she has gained unique insight into the current state of and upcoming trends in the financial industry from a documentation and Legal-Tech perspective. Ng is an author of *The LegalTech Book* (Wiley, 2020).

Frank Pasquale is a professor at the Brooklyn Law School. He is a noted expert on the law of AI, algorithms, and machine learning. He is a prolific and nationally regarded scholar whose work focuses on how information is used across a number of areas. His wide-ranging expertise encompasses the study of the rapidity of technological advances and the unintended consequences of the interaction of privacy law, intellectual property, and antitrust laws. His recent works include *New Laws of Robotics* (Harvard University Press, 2020); *The Oxford Handbook of Ethics of AI* (Oxford University Press, 2020); and *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard University Press, 2015). Pasquale is an affiliate

fellow at Yale University's Information Society Project and a member of the American Law Institute. He is the author of seventy articles and chapters, including "Data-Informed Duties in AI Development," *Columbia Law Review* (2019).

Cristina Poncibò is Professor of Comparative Private Law at the Law Department of the University of Turin, Collegio Carlo Alberto Affiliate and a faculty member at Georgetown Law (Center for Transnational Legal Studies, London). She is also a fellow of the Transatlantic Technology Law Forum (Stanford Law School and Vienna School of Law). Her most recent books include *Contracting and Contract Law in the Age of Artificial Intelligence* (edited; Hart, 2022) and *The Cambridge Handbook of Smart Contracts, Blockchain Technology and Digital Platforms* (edited; Cambridge University Press, 2019). She is the scientific director of the Master's in International Trade Law at the University of Turin, ITC-ILO, in cooperation with Unicitral and Unidroit. In her career, she has been a Marie Curie Intra-European fellow (Université Panthéon-Assas) and a Max Weber fellow (European University Institute).

Wolfgang Schulz is Chair for Media Law and Public Law and is the director of the Hans-Bredow-Institut for Media Research at the University of Hamburg. He is also the director at the Alexander von Humboldt Institute for Internet and Society in Berlin. Schulz is a member of the Committee of Experts on Internet Intermediaries (MSI-NET) and serves on the advisory board of the Law & Technology Centre of Hong Kong University. He is coauthor of *Human Rights and Encryption* (UNESCO, 2016) and *Regulated Self-Regulation as a Form of Modern Government* (Indiana University Press, 2004).

Indra Spiecker genannt Döhmann is Professor of Public Law, Information Law, Environmental Law, and Legal Theory at Goethe University Frankfurt am Main, where she heads the Data Protection Research Unit. She received her PhD from the University of Bonn and completed her habilitation at the University of Osnabrück and received an LLM from Georgetown University. She is member of the Competence Center on IT-Security (KASTEL) at the Karlsruhe Institute of Technology. She is the copublisher of the *European Data Protection Law Journal* and the coeditor of *Computer und Recht* (*Computer and Law*). In 2016, as first lawyer, she was appointed a member of the German Academy of Technical Sciences of the Union of German Academies of Sciences and Humanities. She regularly advises state and private institutions on regulatory matters of digitalization.

Gerald Spindler is Professor of Civil, Commercial, Business, Multimedia, and Telecommunications Law at the University of Göttingen. He studied Law and Economics at Frankfurt am Main. He received research fellowships in Frankfurt, where he was conferred a Comparative Law PhD. His habilitation focused on "Organisational Duties of Companies," for which he received his *venia legendi* for Civil Law, Commercial and Economic Law, International Private Law, Comparative Law, and Labor Law. Spindler is Vice-Chairman of the German Society of Law and Information Science and has advised the German and European legislators on various questions concerning the information society and corporate law.

Eric Tjong Tjin Tai is Professor of Private Law at Tilburg University. He obtained degrees in Computer Science from Delft University of Technology, and in Philosophy and in Law from the University of Amsterdam. Tjong Tjin Tai worked for eight years as a lawyer prior to taking up a position at Tilburg University. His research covers digitalization and private law, service contracts, procedural law, and methodology. He publishes extensively in leading Dutch and European law journals and has authored several books. His current research focuses on the

interaction of IT and human labor in organizations and its consequences for private law, as well as the changes to private law in dealing with new technologies.

Divya Taneja is an associate at Perkins Coie, Seattle, WA, where her practice focuses on technology, media, and intellectual property transactions and counseling across a variety of industries. Taneja has negotiated software license agreements, SaaS agreements, data-sharing agreements, nondisclosure agreements, master services agreements, and other technology-related agreements. She has counseled clients on data privacy and cybersecurity matters, including advising on breach response, drafting privacy policies and procedures, preparing incident response plans, and analyzing compliance with applicable privacy laws and regulations. She is an information privacy professional certified by the International Association of Privacy Professionals.

Susan Y. Tull is a partner in Finnegan's office in Washington, DC, where she specializes in all phases and forums of patent litigation and client counseling. Her patent litigation, appeals, and postgrant proceedings practices focus on technologies related to consumer products, software, AI and machine learning, medical devices, automotive, and other mechanical and electrical systems. Tull has researched and written extensively on patenting AI and software technologies.

Christian Twigg-Flesner is Professor of International Commercial Law at the University of Warwick. Previously, he was associated with the University of Hull, Nottingham Trent University, and Sheffield University. He is a fellow of the European Law Institute, an associate academic fellow of the Honourable Society of the Inner Temple, and coeditor of the *Journal of Consumer Policy*. He has been a senior international fellow at the University of Bayreuth. His research interests are in the areas of international, English, and European commercial, consumer, and contract law, with a particular focus on the implications of digitalization. He has published many articles and book chapters on EU consumer and contract law. His authored or coauthored books include *Rethinking EU Consumer Law* (Routledge, 2017) and *The Europeanisation of Contract Law* (Routledge, 2013). He edited the *Elgar Research Handbook on EU Consumer and Contract Law* (Edward Elgar, 2016) and the *Cambridge Companion to European Union Private Law* (Cambridge University Press, 2010).

Jonathon Watson is a postdoctoral research fellow at King's College London in the National Research Centre on Privacy, Harm Reduction and Adversarial Influence Online. He completed his undergraduate studies in English and German Laws (LLB) at the University of Liverpool and his Dr. iur., LLM at the University of Münster.

D. Sean West is an associate at Perkins Coie, Seattle, WA, where he counsels clients on issues related to intellectual property, commercial transactions, privacy, e-commerce, the Internet of Things, AI, and consumer protection. West works with clients, ranging from start-ups to large public companies, in a variety of industries, including hardware, software, e-commerce, luxury goods, standard bodies, digital media, and the Internet.

Stefan Wrbka is Professor of Business Law at the University of Applied Sciences for Management and Communication, Vienna. Previously, he was an associate professor at Kyushu University. Wrbka earned a Mag. iur. from the University of Vienna, an LLM at Kyushu University, and a PhD (Dr. iur.) from the University of Vienna. Prior to entering academia, he was an in-house counsel at the global headquarters of Red Bull. He is the author or coauthor of more than seventy publications in English, German, and Japanese including *International Business Law: Emerging Fields of Regulation* (Bloomsbury, 2018) and author of

European Consumer to Access Revisited (Cambridge University Press, 2015). He was a coeditor of *The Shifting Meaning of Legal Certainty in Comparative and Transnational Law* (Springer, 2017) and *Flexibility in Modern Business Law: A Comparative Assessment* (Springer, 2016).

Katharina A. Zweig is a professor at the TU Kaiserslautern and head of the Algorithm Accountability Lab. She was a member of the Enquete Commission on Artificial Intelligence for the consultation of the German Parliament (2018–2020); and is on the ITA Advisory Board of the Federal Ministry of Education and Research. She founded a new field of studies called Socioinformatics and her bestselling German book is currently translated and will published under the title "Awkward Intelligence" at the end of the year 2022.

Foreword

There was a time when, as a general rule, lawyers took little interest in technology and, coming from the opposite direction, technologists took little interest in the law. From a lawyering perspective, technology seemed to be neither particularly salient nor significant. Even legal scholars who were interested in setting the law "in context" would not normally highlight technology as an important part of the context. Of course, there were exceptions to this general rule: There were individuals who developed niche interests in particular technologies (such as computing); the imagination of tort lawyers was sometimes stirred by new technologies and their applications (famously so in the case of Warren and Brandeis[1]); and intellectual property lawyers needed to be aware of the changing technological background. However, this was all destined to change and to do so in two disruptive phases.[2]

First, developments in biotechnologies and, at much the same time, in information and communication technologies (ICTs) became salient as a challenge for lawyers – particularly for legislators and regulators.[3] To be sure, some lawyers continued to downplay the significance of these developments,[4] but the challenges were both real and significant. Lawmakers and regulators were expected to make the right interventions at the right time, always balancing the interest in supporting and incentivizing beneficial innovation with the interest in protecting individuals against unacceptable risk, guarding against systemic risk, and respecting a community's fundamental values.[5] While biotechnologies (particularly modern genetics) provoked new questions about respect for human life, human rights, and human dignity,[6] ICTs presented major challenges for national regulators who sought to impose their own rules on the transnational

---

[1] Samuel D. Warren and Louis D. Brandeis, "The Right to Privacy" (1890) 5 *Harvard Law Review* 193.

[2] See Roger Brownsword, *Law, Technology and Society: Re-imagining the Regulatory Environment* (Abingdon: Routledge, 2019) and "Law Disrupted, Law Re-imagined, Law Re-invented" (2019) 1 *Technology and Regulation* 10.

[3] Also, for patent examiners: see Roger Brownsword and Morag Goodwin, *Law and the Technologies of the Twenty-First Century* (Cambridge: Cambridge University Press, 2012), Ch. 1; and Aurora Plomer and Paul Torremans (eds.), *Embryonic Stem Cell Patents: European Law and Ethics* (Oxford: Oxford University Press, 2009).

[4] Compare Frank H. Easterbrook, "Cyberspace and the Law of the Horse" (1996) *University of Chicago Legal Forum* 207.

[5] Compare Brownsword and Goodwin, Law and the Technologies (n. 3) and Roger Brownsword, 'Legal Regulation of Technology: Supporting Innovation, Managing Risk and Respecting Values' in Todd Pittinsky (ed.), *Handbook of Science, Technology and Society* (New York: Cambridge University Press, 2019), 109.

[6] See Deryck Beyleveld and Roger Brownsword, *Human Dignity in Bioethics and Biolaw* (Oxford: Oxford University Press, 2001); and Roger Brownsword, *Rights, Regulation and the Technological Revolution* (Oxford: Oxford University Press, 2008).

online environment that was ushered in by the Internet.[7] Moreover, whether the developments were in biotechnologies or in ICTs, they were coming too thick and too fast for legislators who wanted to enact hard-wired and sustainable legal frameworks. In short, whatever benefits these technologies promised to bring with them, from a legal point of view they were a problem and the challenges that they presented simply could not be ignored.

Secondly, much more recently, developments in AI and machine learning as well as in the use of the blockchain, together with a raft of developments in surveillance, identifying, tracking, and monitoring technologies, have given the legal and regulatory salience of technology a major boost. To some extent, these technological developments are still seen as a problem. Currently, we have lively debates about how to regulate AI, big data, profiling, cryptoassets, facial recognition technologies, deepfakes and so on. Like genetic manipulation and online transactions and content before them, the latest technologies are, so to speak, "out there" as regulatory targets challenging lawyers to find acceptable, effective, and agile ways of governing them. However, the twist – and it is potentially a huge twist – is that we now see these same technologies as tools that can be applied by lawyers and regulators for the smarter performance of legal and regulatory functions, such as the more efficient delivery of legal services.[8] Viewed in this new light, the salience and significance of technology is, distinctively, that it presents lawyers and regulators with fresh options: LawTech (or LegalTech) and RegTech are opportunities. If law is to be fit for purpose, it is not enough that we have the right rules in place and the right kind of people undertaking legal functions; we also need to be thinking about how new tools might be deployed to complement, or even to supplant, our reliance on persons and our reliance on rules and standards.[9]

To see the bigger picture of the relationship between law and technology, we would need to supplement our initial sketch in several respects: We would need to fill in how technologists now view tools for law; we would need to add in the way in which ethicists have related to, and do now relate to, both law and technology (the law/ethics/technology triad); we would need to underline what is being "decentered" as technology moves into the foreground (sidelining public governance, challenging public participation, taking humans and rules out of the loop, and putting the emphasis of governance on ex ante risk management and prevention rather than ex post punishment, correction, and compensation);[10] and we would need to consider whether concepts like justice, trust, authority, and respect continue to be meaningful once governance is turned over to machines, or by machines that are under some degree of human control.[11]

It is against this background of disruption and transformation that this timely Cambridge handbook offers a view of the landscape of lawyering in our digital age, a view that is wide-ranging (covering the impact of AI on the legal profession; alternative dispute resolution; consumers and small claims, and in relation to public law; and the interface between, on the one side, legal ethics and societal values and, on the other, AI) and, importantly, a view that is deep, detailed, and reflective. From the many cues for further discussion given in this impressive

---

[7] Seminally, see David R. Johnson and David Post, 'Law and Borders – The Rise of Law in Cyberspace' (1996) 48 *Stanford Law Review* 1367.

[8] For critical assessment, see, e.g., Karen Yeung and Martin Lodge (eds.), *Algorithmic Regulation* (Oxford: Oxford University Press, 2019); and Simon Deakin and Christopher Markou (eds.), *Is Law Computable?* (Oxford: Hart, 2020).

[9] Compare Brownsword, Law, Technology and Society (n. 2) and Roger Brownsword, *Law 3.0: Rules, Regulation and Technology* (Abingdon: Routledge, 2020).

[10] For the decentering of law, see Roger Brownsword, *Rethinking Law, Regulation and Technology* (Cheltenham: Edward Elgar, 2022), Ch. 1.

[11] See, further, Roger Brownsword and Han Somsen, 'Law, Innovation and Technology: Fast Forward to 2021' (2021) 13 *Law, Innovation and Technology* 1.

book, let me pick up on one of the concluding remarks which is to the effect that, in addition to disrupting the practice of law, AI "will have a major impact on legal education and legal ethics" – how right that surely is.

To start with the disruption of law schools, there have been numerous silver linings stemming from COVID. It might have led to "greener" policies and practices, and it has acted as an accelerator for the use of digital technologies in the delivery of legal education and in the undertaking of legal research. However, the disruption that is now contemplated is not so much in *how* we teach or research the law, it is in *what* we teach and *what* we research (as well as with whom we research).

In the law schools, it is axiomatic that our mission is to train students "to think like lawyers." When lawyers had no interest in technology, thinking like a lawyer was inward-looking, doctrinal, and guided by general principles. By the time that lawyers saw technology as a challenge, thinking like a lawyer needed to become more "regulatory," more policy-orientated, and more outward-looking (with a view to learning from economics, sociology, and philosophy), but, generally, the law schools persisted with their habitual doctrinal approach and treated regulatory scholarship as marginal to legal education. Now that we are in the age of digital law, it makes no sense to resist thinking that is both regulatory and interested in developing technical solutions. To think like a lawyer now has to be a three-dimensional exercise (engaging with general legal principles, with regulatory approaches, and with technological solutions to regulatory challenges). Moreover, thinking like a lawyer has to be guided by an overarching understanding of the importance of the critical infrastructures on which human communities are predicated – an appreciation of which holds the key to whether the modality, the process, and the substance of governance is legitimate.[12] Precisely how law schools should rewrite their curriculum to reflect this rethinking of what it is to think like a lawyer, of what it is to have the relevant legal skills, is now the big question.[13]

No less a challenge is presented by the future of legal research. Once upon a time, the agenda for legal research was almost entirely doctrinal. Researchers took the coherence and integrity of legal doctrine (and, similarly, of adjudication) as the benchmark and had no difficulty in uncovering the many tensions and contradictions that were often flagged by the leading cases. For regulatory scholarship, the agenda was rather different, but it was dominated by instrumental concerns, by identifying which regulatory interventions worked and which did not (and why). As I have said, this kind of legal scholarship invited collaboration with other disciplines, but not usually with science and technology. Today, the agenda is moving on. In just the way that the handbook proposes and, indeed, exemplifies, lawyers need to figure out how to engage with the technological dimension of legal thinking and this particular modality of governance. Whether or not the legal scholars of tomorrow also need to become data scientists or some other kind of technologist is a moot question but, in any event, if they are to engage with governance by technologies, lawyers will surely need to be involved in a much wider range of interdisciplinary collaborations. Crucially, lawyers need to be in the vanguard in developing our intelligence about how best to discharge legal and regulatory functions. To be sure, as questions about AI

[12] See further Brownsword, *Law, Technology and Society* (n. 2), and *Rethinking Law* (n. 10).
[13] See, e.g., Brownsword, *Rethinking Law* (n. 10), Chs. 14 and 15; Mark Fenwick, Wulf A. Kaal, and Erik P. M. Vermeulen, 'Legal Education in the Blockchain Revolution' (2017) 20 *Vanderbilt Journal of Entertainment and Technology Law* 351; Julian Webb, 'Information Technology and the Future of Legal Education: A Provocation' (2019) 7 *Griffith Journal of Law and Human Dignity* 72; Mark Findlay, *Globalisation, Populism, Pandemics and the Law* (Cheltenham: Edward Elgar, 2021) esp. at 146; and William N. Lucy, 'Law School 2061' (2021) 84 *Modern Law Review* (2022) 85 Mod. L. Rev. 1468.

soon reveal, digital law brings with it a good deal of uncertainty – we know that there are things that we do not know about governance by technologies and we know that there are major challenges in operating under conditions of uncertainty but, before we despair, we should at least gather up and synthesize the things that we do know. Lacking a hub that gathers such intelligence, nationally and internationally, is a major weakness in our institutional arrangements.[14]

These comments take me to a few words about ethics. The story of the relationship between law and ethics is a bit like that of law and technology. Initially, law has no interest in ethics as an external reference point; ethics at this stage has been internalized in various doctrines (such as general principles that require good faith or reasonable endeavors, or that provide relief against unconscionable conduct). However, at the same time that law comes to see technology as a challenge, it begins to take an interest in external (particularly professional medical and health care) ethics but, in order to improve the chances of regulatory interventions being treated as acceptable, a broader range of community ethics needs to be taken into account. In the present phase, when law sees technology as an opportunity, we find that some regulators are very ready to be guided by expert ethics groups (the case of AI applications in the EU is a striking example) but we also find that technologists and tech enterprises are keen to shape ethical thinking. The external ethification of law, the relationship between tech-generated codes and the law, and the relationship between local ethics and cosmopolitan standards are urgent topics for research.[15] In the bigger picture, the drivers behind LawTech and RegTech are efficiency, convenience, and workability and, while these are relevant considerations, it is important that ethics operates as a counterpoint to a discourse that is largely instrumental.[16]

Finally, I am confident that one thing that reviewers will say is that all lawyers should read this handbook. Not only is the handbook a valuable guide to where we stand in the evolution of the relationship between law and technology, it is an essential preparation for the turbulence that will be experienced as new technologies and novel applications create new tools that insinuate themselves into every corner of the practice of law.

Roger Brownsword

Professor Roger Brownsword is a professor of Law at King's College London. He was the founding director of King's College Centre for Technology, Ethics, and Law in Society in 2007. He is the founding general editor of *Law, Innovation and Technology* as well as being on the editorial board of the *Modern Law Review*, the *International Journal of Law and Information Technology*, and the *Journal of Law and the Biosciences*.

---

[14] See, further, Brownsword, *Rethinking Law* (n. 10), Chs. 10 and 11.

[15] On the last-mentioned, see Roger Brownsword, 'Migrants, State Responsibilities, and Human Dignity' (2021) 34 *Ratio Juris* 6.

[16] Compare, e.g., Ethan Katsh and Orna Rabinovich-Einy, *Digital Justice* (Oxford: Oxford University Press, 2017); and Benjamin H. Barton and Stephanos Bibas, *Justice Rebooted* (New York: Encounter Books, 2017).

Preface

Due to the acceleration of technology, books dealing with issues involved with the subject quickly become obsolete. This book attempts to avoid this outcome by discussing the current and future states of artificial intelligence (AI). It provides commentary on existing AI systems, focusing on its enhancing and disruptive effects on the current state of law. It also provides insight regarding the function of law in a future of advanced AI or superintelligence. Finally, it will discuss the dimensions of human intelligence and artificial intelligence from both legal, societal, and ethical perspectives.

This book comprises a collection of chapters focused on the role of law and ethics in the development and application of AI. It should be noted that all the editors equally contributed to this book project. The topics selected for the book seek to present a range of perspectives on the rise of AI and to theorize about how best to plan for the future. The book is broad in scope, discussing numerous areas of laws and uses of AI. Some of the broad issues discussed include definition and operation of AI systems, benefits and pitfalls of AI, preventing AI bias, security and privacy of personal data, issues posed by physical manifestations of AI, risks and threats of the interconnectivity of AI to the Internet of Things, developing public policy relating to AI, benefits and threats of the future development of "superintelligence," ethical standards and guidelines for AI, and whether AI should be granted personhood. More specific to practitioners is the discussion of AI's impact on the practice of law, on the negotiation and performance of contracts, as well as on corporate governance, and tort theory's application to AI from European and American perspectives. Other areas covered include application of agency law to AI agents, liability for harm caused by AI systems (autonomous vehicles), AI's interrelationship with competition law, use and regulation of robo-judges, patentability of AI, and AI as inventor or creator.

We are grateful for the work and patience of our numerous contributors. Due to the pandemic the project was postponed several times. It is only through the honoring of their commitments that this book was made possible. We are also indebted to the University of Torino, University of Florida, and Lyon Catholic University for their support. Special thanks to the commissioning and editorial staff at Cambridge University Press, especially Matt Gallaway.

AI: Development and Trends

Artificial Intelligence

The Promise of Disruption

Larry A. DiMatteo

1.1 INTRODUCTION

Disruption – societal and economic – has been a part of humankind from the beginning of time. One example is the transition from a mostly agrarian economy to the industrial age toward the end of the nineteenth century. Over time stalwart career paths were made obsolete. The need for blacksmiths gradually diminished, while the demand for welders arose. The technological age is just another example although the disruption of employable skills and ways of doing business has been amplified due to the recent acceleration of technological development. Advanced artificial intelligence (AI) or superintelligence promises much greater disruption. Although disruption of the status quo is viewed in a derogatory sense, mainly by those anchored in the status quo, taken from the broader view of the betterment of humankind disruption has been a positive force in the macro sense. However, advancement or disruption does not come without costs – the industrial age put the world on the path toward the existential threat of climate change. Nevertheless, this was not an inevitable pathway. If the wealthier nations had the political will, aided by technological breakthroughs, then the environment crisis could have been avoided or diminished.

The world is close to reaching another inflection point: the so-called existential threat of superintelligence with the potential of replacing human control and decision-making with its creation. Before that point, AI and other technologies have caused major disruption in the economy and employment, and this disruption will only accelerate in the future. The pivotal issue is not whether advanced AI is preventable. It is not. Even if it can be delayed there are strong utilitarian and deontological arguments that favor the encouragement of AI development. The focus should be on mitigating the negative effects of disruption and using smart design to prevent AI from ever becoming an existential threat to humankind. Professor John O. McGinnis makes an eloquent argument along these lines in his chapter entitled: "The Folly of Regulating against AI's Existential Threat" (Chapter 27).

Artificial intelligence is currently used in many areas of society – economic, big data, and government activities. It has shown its many benefits in medicine, industry, consumer marketing, and so forth. The acceleration of technology has made it obvious that AI will continue to get smarter with greater abilities to make decisions previously made by humans. The AI of the future will be characterized by greater degrees of autonomy in searching big data,

accelerated machine learning, and physical robots. This brings up the issue of – despite the benefits of AI – what threats does AI pose to society and democratic institutions?[1]

Movers in society, law, design, and technology need to work together to avoid the pitfalls of future advanced AI or superintelligence. Automation has been a core factor in creating more efficient and wealthier economies. However, automation coupled with AI has led to the fear that AI decision-making, which lacks human involvement, will threaten our way of life. A truly autonomous AI system may make decisions not wanted or expected. Some have suggested that the fear of AI as an existential threat is misplaced because it will never be able to replicate the human mind. Nobel Laureate Daniel Kahneman notes that AI is only able to master System 2 thinking, which entails deliberative, rational thinking, but not System 1 thinking, consisting of intuitive and creative thinking, which will remain the domain of human beings.[2] This may be true but there is still the threat that unregulated, unmonitored autonomous systems will sometime in the future coopt the ability of humans to intervene with System 1 thinking.

This chapter sketches out the many issues related to advanced AI and its governance. It reviews the legal and ethical dimensions of that governance. Due to the acceleration of technology, the legal response will lag behind, but at some point, law will have to become forward-thinking in order to anticipate and prevent the dangers that future AI presents. For now, a few rules of thumb can be noted. AI decision-making must not be totally autonomous and must allow for human intervention. Current legal and ethical concepts, such as agency, autonomy, fairness, contract, property, intellectual property (IP), and trust, will be flexible enough to regulate abuse in the early stages of development. At some point in time specialized regulation will need to be created. Specialized regulation will include the use of technological design to ensure compliance with the law. At the ethical level, society should make normative assessments of when the pure efficiency of technological advancement is outweighed by the quality of human life and community. In the end, any solutions to future dangers presented by superintelligence must be an interdisciplinarity process that includes the input of policy makers, technologists, ethicists, and lawyers.[3]

It is clear that human beings will have an ongoing role to play in supervising AI from ethical and legal perspectives. AI will need to be monitored to prevent it from overshadowing the human element. AI will know the world differently than the way humans do; it does not possess first-person experiences that humans have developed over millennia through evolutionary processes. To lose that element would be to lose a vital piece of our humanity.

This book takes a broad view of the current and future uses of AI. It is structured along a number of topical areas. The areas chosen for study show the broad impact of AI, but remain only a selective sampling of the areas, now and in the future, that AI will impact. This introductory chapter discusses AI from a broader societal view through the perspectives of law, ethics, and public policy.

---

[1] This future threat was symbolized by the computer HAL in Stanley Kubrick's 1968 movie *2001: A Space Odyssey*. HAL represents humans' greatest achievement, which evolves to threaten the destruction of its human overlords.

[2] Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus & Giroux, 2011). For a fuller discussion of the application of Kahneman's theory of thinking see Joshua Davis, "AI, Ethics and Law: A Possible Way Forward," Chapter 21, pp. 306–307.

[3] "Interdisciplinary teams are necessary for AI and application design to bring together technical developers with experts who can account for the societal, ethical and economic impacts of the AI system under design." HUB4NGI, "Responsible AI – Key Themes, Concerns & Recommendations for European Research and Innovation" (June 2018), www.ngi.eu. Permission of Steve Taylor (S.J.Taylor@soton.ac.uk).

The most positive understanding of the value of AI is to see it as a public good: "One area of great optimism about AI and machine learning is their potential to improve people's lives by helping to solve some of the world's greatest challenges and inefficiencies."[4] However, many things introduced to advance the public good work for and against that good. As such, if good is used as the basis of power the illicit use of that power is likely to follow. Something promoted for the public good will still need to be regulated. The key is that regulation must be focused on the overreach of the power that is AI, while not retarding its development. The right kind of regulation works hand in hand with the expansive use of AI. The more any misuse of AI can be prevented (or punished) the greater will be the trust in its development as a safe means to cure the many problems the world faces (e.g., climate change, poverty, equal opportunity, pandemics, and scarcity of resources).

Despite the many benefits that AI promises its development must be placed within a broader landscape of public policy. Automation, for example, is disruptive of the current employment needs of companies by making current skills sets obsolete. A forward-looking industrial policy, which includes the transitioning of workers into the skill sets created by AI, will need to be created. Foreseeable disruption without planning and mitigation will unleash the demons of human nature. The consequences of disruption caused by AI can be lessened by public policies and programs that sit outside of AI innovation.

The ethical implications of replacing human decision-making with AI decision-making must also be forward thinking. There are elements that need to be incorporated into the process of AI development – transparency, human intervention, and skills training. In the area of transparency, understanding the process of AI decision-making is vital to any interests that are impacted by such decisions: "Transparency concerns focus not only on the data and algorithms involved, but also on the potential to have some form of explanation for any AI-based determination." However, the ability to understand advanced AI systems, as well as predicting their behavior, is problematic. Therefore, the second element of ethical AI requires the design of AI that allows for human intervention to monitor and overturn AI decision-making. Finally, technical skills must be accompanied by ethical and legal skills in the design and use of AI. Technological progress of what is possible needs to be done in a framework of "putting good intentions into practice by doing the technical work needed to prevent unacceptable outcomes."[5]

There are many questions being debated about the use of advanced AI, both specific and broad in nature. More specifically, what are the implications that AI systems pose for privacy, security, and protection of personal and sensitive data? What are the ethical implications of AI's use in dealing with consumers? Should AI be granted personhood? More broadly, how should ethical standards and guidelines be developed for AI? How should public policy be constructed relating to AI? What are the benefits and threats of the future development of superintelligence?

This chapter will analyze current soft law instruments and literature to determine the ways society may minimize the risks of superintelligence becoming an existential threat to humanity.[6] Types of regulations to be considered include targeted statutory law, self-regulation, and

---

[4] National Science and Technology Council, "Preparing for the Future of Artificial Intelligence" (October 2016), 2.

[5] National Science and Technology Council, "Preparing for the Future," 3.

[6] The idea of a robot takeover has been the grist for many sci-fi movies. Some of the doomsday scenarios pose such questions as, "What if one day machines decided that humans were just a waste of resources and started a robocalypse? Or, will artificial general intelligence be humanity's last invention?" See Doomsday Now, "Robot Takeover," https://doomsdaynow.com/robot-takeover/.

standardization.[7] In some ways an analogy can be drawn between the development of advanced AI and the cloning of human beings. Despite the perceived benefits of cloning, such as growing human organ replacements, cloning of human beings has been universally condemned for medical, safety, and ethical reasons. The biggest concern is that it would lead to the creation of "better human beings" violating principles of equality and human dignity. In the same way, the growth of superintelligence threatens the autonomy and dignity of human beings. This chapter will explore the ways that can be used to prevent this from happening.

Section 1.2 explores the nature of law – how it evolves and how this evolution lags behind real-world developments. This lag is generally beneficial because it allows new things to develop more quickly and incentivizes innovation. However, in the age of the acceleration of technology[8] legal gradualism poses a problem in managing the development of advanced AI and fortuitously mapping out impermissible AI systems and applications. Section 1.3 will review the types of principles offered by soft law instruments to encourage the responsible development and use of AI. Section 1.4 discusses the genesis of a European approach to the development of trustworthy AI. Section 1.5 reviews the coverage of the book.

1.2 NATURE OF LAW

The relationship between law and society can be framed a binary one. On the one hand, law needs to respond to developments in society or face becoming obsolete. In this way law is purely reactive in nature. On the other hand, law can be a positive force in the development of society by placing normative limits on social development or what Karl Llewellyn called the "marking off of the impermissible."[9] In this way, law plays a proactive role in shaping how society evolves.

The evolution of law is reactive in nature resulting in a lag between real-world developments and their regulation. The virtue of "lag" is that premature regulation of something new may stifle its development and discourage innovation. The history of the Internet is an example of determining when the best time is to regulate a new technology. There were two points of view regarding the regulation of the Internet – the libertarian view that it should not be regulated in order to allow for it to continue to develop unimpeded and the traditionalist view that the newness of the Internet and unknown dangers posed by such technology required targeted law to prevent abuse. In this case, the libertarian view won out leading to the central role the Internet now plays in daily life. In recent years serious consideration has been given to enact laws, such as the EU General Data Protection Regulation (GDPR), to manage the threat that social media companies and big data, enabled by the Internet, present to human autonomy and dignity. This is the issue now presented by AI: Should it be regulated or not and, if so, when

---

[7] Standards provide requirements, specifications, and guidelines to be applied to ensure that AI meets its technical and ethical objectives. Standards can address issues in the areas of software engineering (security, monitoring), performance (accuracy, reliability, scalability), safety (control systems, regulatory compliance), interoperability (data, interfaces), security (confidentiality, cybersecurity), privacy (control of information, transmission), traceability (testing, curation of data), and so forth. See National Science and Technology Council, "The National Artificial Intelligence Research and Development Strategic Plan" (October 2016), 32–33.

[8] Thomas Friedman states that: "Technology is now accelerating at a pace the average human cannot keep up with." MIT News, "Thomas Friedman Examines Impact of Global Accelerations" (October 2, 2018), https://news.mit.edu/ 2018/thomas-friedman-impact-global-accelerations-1003#:~:text=%E2%80%9CTechnology%20is%20now%20accelerat ing%20at%20a%20pace%20the,added%2C%20emphasizing%20a%20key%20theme%20of%20his%20talk.   See   also Thomas Friedman, *Thank You for Being Late: An Optimist's Guide to Thriving in the Age of Accelerations* (New York: Farrar, Straus & Giroux, 2016).

[9] Karl N. Llewellyn, "Book Review," *Harvard Law Review* 52 (1939): 700, 704.

should it be regulated and how? In this case, the threats to society are on a larger scale in that the creation of better autonomous systems that will lead to a major shift from human decision-making to machine decision-making. The lure of autonomous AI decision-making is that it is more accurate and efficient. The era of big data makes automated processes a necessity.

The reactive nature of law through the ages has generally been a positive feature mainly because of the gradual nature of change. Today, the acceleration and complexity of technology renders law enfeebled in the face of modernity. This presents the problem that if law continues to lag behind the technological advancement of superintelligence (independent decision-making with the ability of human intervention) and super-superintelligence (the loss of the ability of humans to intervene) any regulation will prove to be futile. This is seen in what has been called the alignment problem[10] of advanced AI in which the autonomous system something that it "believes" is in the best interest of its human benefactor and instead the decision is not the one that the human being would have made. Stated differently, AI makes a value judgment that is not aligned with the values or expectations of the human parties. This is akin to the agency problem found in corporate law where director–officer–employee interests may diverge with the interests of the corporation and its shareholders.[11]

Law to be effective will have to be proactive. An analogy can be seen in evolutionary biology where Stephen Jay Gould contested evolutionary theory as a slow, gradual process by showing that the fossil records indicate times of evolutionary "jumps" or "punctuated equilibria."[12] In the area of AI, the law will at some point of development of AI need to jump ahead in order to prevent future dangers from occurring. Hopefully, the "angst of futuristic surrender to an AI and robotically controlled world" will provide the motivation for proactive regulation.[13]

The regulation of future, unknown technological developments seems to be an impossible task, but in fact a plausible regulatory framework can be envisioned to regulate perceived future threats. The precautionary principle used in environmental protection is a case in point. Principle 15 of the United Nations' Rio Declaration on Environment and Development states: "In order to protect the environment, the precautionary approach shall be widely applied by States according to their capabilities. Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation."[14] Stated more simply, the precautionary principle means that an action should not be taken if the consequences are uncertain and potentially dangerous. Thus, if the consequences of a technological advancement pose potential dangers to human dignity, human rights, or democratic processes then it should be prohibited despite its perceived benefits. This is a rejection of the notion that the benefit of AI is an unassailable truth.

---

[10] Peter McBurney and Simon Parsons, "Talking about Doing," in Katie Atkinson, Henry Prakken, and Adam Wyner (eds.), *From Knowledge Representation to Argumentation in AI: Law and Policy Making* (London: College Publications, 2013), 151–166.

[11] See Patrick McColgan, "Agency Theory and Corporate Governance: A Review of the Literature from a UK Perspective" (May 22, 2001), https://pdfs.semanticscholar.org/79c5/2954af851c95a27cb1fb702c23feaae86ca1.pdf.

[12] Stephen Jay Gould and Niles Eldredge, "Punctuated Equilibria: The Tempo and Mode of Evolution Reconsidered," *Paleobiology* 3 (1977): 115–151.

[13] This danger is known as known as "'singularity,' whereby superintelligent machines take over and permanently alter human existence through enslavement or eradication." Mike Thomas, "The Future of Artificial Intelligence," https://builtin.com/artificial-intelligence/artificial-intelligence-future (updated April 20, 2020).

[14] Report of the United Nations Conference on Environment and Development, A/CONF.151/26 (Vol. I) August 12, 1992, www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_CONF.151_26_Vol.I_Declaration.pdf.

Big data and analytics have shown that the harvesting of personal data can produce high profits. This type of monetary incentive will result in the amoral exploitation of AI. An analogy is seen in blockchain technology, which provides a secure, efficient, and anonymous vehicle for transferring information, but in the wrong hands it can be used to illegally launder money. Just as clandestine laboratories may seek to illegally clone a human being, incredulous enterprises may seek to develop types of AI and applications prohibited by future law. The pervasiveness and depth of regulation and monitoring will be pivotal in stemming such illicit activities. The rest of the chapter will analyze a basket of regulations that can be used to prevent the exploitation of AI.

1.3 RESPONSIBLE AI

Dianna Wallis sees the speed of technological development and the complexity of the issues it presents as a call to arms. She asserts that "the sooner we start as a society discussing the issues now presented by advanced technologies, such as AI and superintelligence, the more it is likely that national and international legal systems can develop a holistic approach to the appropriate use and ethical safeguards related to such technologies."[15] This approach requires law and policy makers to be proactive. Instead of waiting until AI poses a threat to democracy and endangers society, "AI needs to be guided by a deliberative political process, to determine how fast and how far such technology should go."[16] The impact of AI programs on democratic processes has been seen in recent elections and noted by the Council of Europe: "AI-based technologies used in online media can contribute to advancing misinformation and hate speech, create 'echo chambers' and 'filter bubbles' which lead individuals into a state of intellectual isolation."[17]

1.3.1 *Landscape of AI, Society, Law, and Ethics*

AI is already in use in many sectors of society touching large companies, government operations, and the consumer marketplace. Figure 1.1 shows one view of an increasingly complex relationship between AI and its stakeholders. It focuses on six themes that need to be considered in creating responsible AI: (1) regulation and control, (2) transparency, (3) responsibility, (4) design, (5) ethics, and (6) socioeconomic impact.[18]

In the area of ethics, society must develop a framework of applied ethics for AI. This includes the selection of existing ethical norms and new norms that adhere to the use of AI. Transparency relates directly to information and education. AI or the humans implementing AI systems must be required to disclose the nature of the processes being used, the personal information being processed, and how decisions are made. Regulation and control require that humans remain in control of autonomous systems including the ability to intervene to change an AI decision. The

---

[15] Diana Wallis, "Visions of the Future," in Larry DiMatteo, Michel Cannarsa, and Cristina Poncibò (eds.), *Cambridge Handbook on Smart Contracts, Blockchain Technology and Digital Platforms* (New York: Cambridge University Press, 2020), 363–364.

[16] Wallis, "Visions of the Future," 368.

[17] Council of Europe, Report of Committee on Political Affairs and Democracy, "Need for Democratic Governance of Artificial Intelligence," Doc. 15150 (September 24, 2020), 9–10.

[18] See also, Jessica Fjeld, Nele Achten, Hannah Hilligoss, Adam Nagy, and Madhulika Srikumar, "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI," Berkman Klein Center Research Paper No. 2020-1 (January 15, 2020), https://cyber.harvard.edu/publication/2020/principled-ai. This study provides a meta-analysis of thirty-six AI principles documents and finds eight prominent themes in order of emphasis: privacy, accountability, safety and security, transparency and explainability, fairness and nondiscrimination, human control of technology, professional responsibility, and promotion of human values.
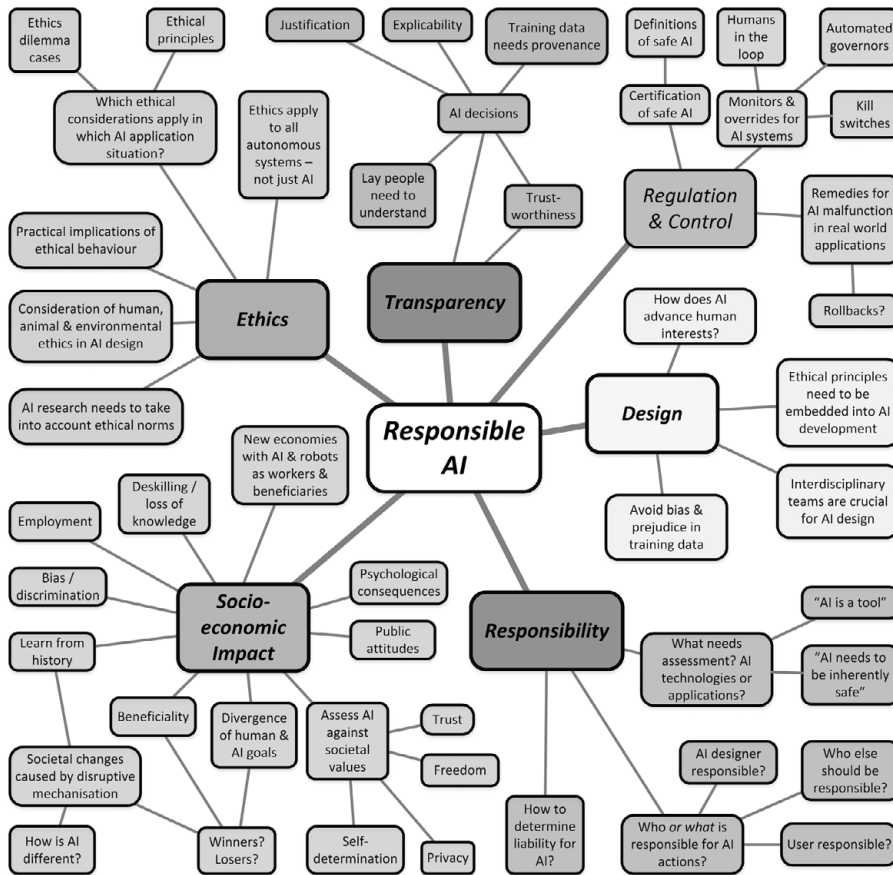
FIGURE 1.1 "Responsible AI" (six main themes)[19]

larger issue is a determination of what is safe AI and the areas where the use of AI is considered inappropriate. This type of assessment must be done through interdisciplinary dialogue as the issues involved cut across the fields of law, computer science, ethics, and technology. It manifests the urgency to turn future conversations on these questions into a "liquid network,"[20] an interdisciplinary space expanding and generating a reliable flow of knowledge.[21]

Design may consist of the use of technology to manage and monitor itself. AI systems must be designed to act ethically and ensure personal information is protected by design. Also, great precaution must be taken so that AI does not replicate the biases of its human programmers. Responsibility is the determination of which of the stakeholders – programmers, creators, owner-users – should be allocated liability if the AI system fails or causes harm.[22] Finally, due diligence

[19] HUB4NGI, "Responsible AI."
[20] Steven Johnson, *Where Good Ideas Come From* (New York: Penguin Group, 2010), 45.
[21] Barbara Pasa and Larry A. DiMatteo, "Observations on the Impact of Technology on Contract Law," in Larry A. DiMatteo, Michel Cannarsa, and Cristina Poncibò (eds.), *Cambridge Handbook on Smart Contracts, Blockchain Technology and Digital Platforms* (New York: Cambridge University Press, 2020), 338, 347.
[22] An example would be how does current product liability law apply to AI systems? See Irina Carnat, "The Notion of Defectiveness Applied to Autonomous Vehicles: The Need for New Liability Bases for Artificial Intelligence," *Trento Student Law Review* 2 (2020): 15 (notes five levels of vehicle autonomy; concludes that the American risk-utility and European consumer expectation approaches to product liability is ill-suited to AI; a better approach would include the development of harmonized technical standards to be applied in the development of autonomous vehicles).

on the socioeconomic impact of AI should be undertaken before its creation and throughout its life cycle. The question to be asked is whether the benefits of AI outweigh its socioeconomic disruptive impact and negative effects on human well-being, which range over concerns of trust, privacy, democratic values, psychological impact in and outside the workplace, and human rights. In sum, just because AI can do something doesn't mean it should be allowed to.

1.3.2 *What Is Wrong with Existing Law?*

The decision not to overly regulate the Internet with specialized bodies of rules proved to be the right decision at least in the beginning of the era of information. Existing legal constructs proved flexible enough to deal with the issues presented. Traditional contract, tort, and intellectual property concepts proved amazingly malleable in controlling internet abuses. For example, the oldest of common law causes of action – trespass – has been used to litigate the improper encroachment on a party's bandwidth. Roger Brownsword refers this ability to fit novel real-world change to existing legal frameworks as the coherentist approach where the fit is a product of manipulation:

> Faced with new technologies, the coherentist tendency is to apply the existing legal framework (the traditional template) to innovations that bear on transactions, or to try to accommodate novel forms of contracting within the existing categories. We need only recall Lord Wilberforce's much-cited catalogue of the heroic efforts made by the courts – confronted by modern forms of transport, various kinds of automation, and novel business practices – to force "the facts to fit uneasily into the marked slots of offer, acceptance and consideration" or whatever other traditional categories of the law of contract might be applicable.[23]

In the words of Brownsword, the regulatory-instrumentalist approach provides an alternative approach. It looks at policy not doctrine as they pertain to particular communities and fundamental values. The difference in approaches is shown in this question: "Even if transactions are largely automated, are there not still Rule of Law concerns implying that there will be some limits on the permitted use and characteristics of [the technology]?"[24] In the end, a combination of both approaches may be needed. Existing legal constructs should be retained and applied to new technologies, but that application should be based on an overt discussion of how best and for what purposes should those constructs be applied. Regulatory instrumentalism will be needed when the peripheral use of existing constructs meets their limitations or borders and more specialized new laws will be needed to align a new technology, such as advanced AI, to community values. This is the point when the benefits of technocracy and efficiency must yield to core values of democracy and human dignity.

1.3.3 *Escaping the Law*

With the advancement of AI and machine learning, a paradigmatic shift, sometime in the future, may be in store, where code will be seen as having the effect of law ("code

---

[23] Roger Brownsword, "Smart Transactional Technologies, Legal Disruption and the Case for Network Contracts," in Larry A. DiMatteo, Michel Cannarsa, and Cristina Poncibò (eds.), *Cambridge Handbook on Smart Contracts, Blockchain Technology and Digital Platforms* (New York: Cambridge University Press, 2020), 313, 322, quoting Lord Wilberforce in *New Zealand Shipping Co Ltd.* v *A. M. Satterthwaite and Co Ltd.*: The Eurymedon [1975] AC 154, 167.

[24] Brownsword, "Smart Transactional Technologies," 332.

is law").[25] Lawrence Lessig has argued that coders and software programmers, by making a choice about the working and structure of IT networks and the applications that run on them, create the rules under which the systems are governed. The coders therefore act as quasi-legislators. In other words, "code is law" is a form of private sector regulation whereby technology is used to enforce the governing rules.[26] This may be true as a technological fait accompli, but it may not be legal or ethically just. For example, an illegal term cannot be made legal simply by placing a contract on a blockchain. Even though the term will be self-executing, and the contracting parties may have little recourse, these characteristics do not magically make the term legal.

The above example is seen as an attempt to escape the law and the court system. The future of AI will provide a similar scenario but in a more potent form. Will democratic and communal values be lost to technological decision-making? Democracy is not the most efficient of governing systems often infected by waste and corruption. In order to prevent such infections, it may be tempting to turn over governmental activities to AI that will be able to make incorruptible and efficient decisions. In this scenario, AI will rise above the law. This would lead to a diminishment in human value and dignity. AI systems lack the human empathy and judgment so vital to human governance. As stated earlier, the threat is that the advancement of AI may proceed to a point where human intervention is no longer possible. In the short term, responsible AI must be developed and monitored in order to protect basic human values. In the long term, further development of truly autonomous AI systems may have to be prohibited.[27]

1.4 REGULATING AI: AREAS OF CONCERN

The depth of the legal, ethical, and policy literature on AI is enormous. For this reason, this section will focus on the initiatives undertaken by the European Union and the Council of Europe. Many of the issues discussed above are recognized in these documents and some solutions are offered. In many cases, however, no specific solution is offered but a pathway to future regulation is given. In 2017, the European Economic and Social Committee (EESC) identified that the most important AI "societal impact domains include: safety; ethics; laws and regulation; democracy; transparency; privacy; work; education and equality."[28] The European Union and the Council of Europe have recognized the need to work toward a regulatory-ethical scheme to deal with future advances in AI. The European Commission, in June 2018, established the High-Level Expert Group on Artificial

---

[25] Jia Wang and Lei Chen, "Regulating Smart Contracts and Digital Perspectives: A Chinese Perspective," in Larry A. DiMatteo, Michel Cannarsa, and Cristina Poncibò (eds.), *Cambridge Handbook on Smart Contracts, Blockchain Technology and Digital Platforms* (New York: Cambridge University Press, 2020), 183, 194.

[26] Lawrence Lessig, *Code and Other Laws of Cyberspace* (New York: Basic Books, 1999).

[27] See Dirk Helbing et al., "Will Democracy Survive Big Data and Artificial Intelligence?" in Dirk Helbing (ed.), *Towards Digital Enlightenment* (London: Springer, 2019), 73–98; Steven Livingston and Matthias Risse, "The Future Impact of Artificial Intelligence on Humans and Human Rights," *Ethics & International Affairs* 33 (2019): 141–158.

[28] EESC Opinion on AI and society (INT/806, 2017). It should be noted that the EU's major concern in the beginning was to encourage the development of AI. In communications of April 25, 2018 and December 7, 2018, the European Commission set out its vision for AI, which supports "ethical, secure and cutting-edge AI made in Europe." "Three pillars underpin the Commission's vision: (i) increasing public and private investments in AI to boost its uptake, (ii) preparing for socio-economic changes, and (iii) ensuring an appropriate ethical and legal framework to strengthen European values." COM(2018)237 and COM(2018)795.

Intelligence,[29] which began work on "Ethics Guidelines on Trustworthy AI" (*Trustworthy AI*).[30] Subsequently, in 2020, the Council of Europe (COE) Ad hoc Committee on Artificial Intelligence (CAHAI) published "Towards a Regulation of AI Systems" (*Towards Regulation*).[31] These documents discuss the "impact of AI on human rights, democracy and rule of law; development of soft law documents and other ethical-legal frameworks; and drafting of principles and providing key regulatory guidelines for a future legal framework."[32] These documents will be discussed below. The following three sections explore the principles needed to guide the future regulation of AI, AI's threats to human rights, and the elements of trustworthy AI.

1.4.1 *Future Regulation of AI*

Towards Regulation ferrets out a number of ethical themes and related issues:

1. Justice is mainly expressed in terms of fairness and prevention (or mitigation) of algorithmic biases that can lead to discrimination; fair access to the benefits of AI (designing AI systems especially when compiling the training datasets).
2. Nonmaleficence and privacy: misuse via cyberwarfare and malicious hacking (privacy by design frameworks).
3. Responsibility and accountability: includes AI developers, designers, and the entire industry sector.
4. Beneficence: AI should benefit "everyone," "humanity," and "society at large."
5. Freedom and autonomy: freedom from technological experimentation, manipulation, or surveillance (pursuing transparent and explainable AI, raising AI literacy, ensuring informed consent).[33]
6. Trustworthiness: control should not be delegated to AI (processes to monitor and evaluate the integrity of AI systems).
7. Dignity: prerogative of humans but not of robots; protection and promotion of human rights; not just data subjects but human subjects.[34]

Towards Regulation incorporates an Israeli Study[35] and Ethics Report.[36] The Study notes that due to the increasing complexity of AI systems "it is difficult to anticipate and validate their behavior in advance."[37] The Ethics Report enunciates six ethical principles central to creating public policy relating to AI:

---

[29] https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence.
[30] European Commission, "Ethics Guidelines on Trustworthy AI" (First Draft, December 2018), April 8, 2019, https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.
[31] COE CAHAI, "Towards Regulation of AI Systems," DGI (2020), 16.
[32] COE CAHAI, "Towards Regulation," 7.
[33] An analogy can be drawn to the GDPR's "right to be forgotten."
[34] COE CAHAI, "Towards Regulation," 53–55.
[35] Isaac Ben-Israel, Eviatar Matania, and Leehe Friedman, "Harnessing Innovation: Israeli Perspectives on AI Ethics and Governance," Report for CAHAI, https://sectech.tau.ac.il/sites/sectech.tau.ac.il/files/CAHAI%20-%20Israeli%20Chapter.pdf; COE CAHAI, "Towards Regulation," 120.
[36] *The National Initiative for Secured Intelligent Systems to Empower the National Security and Techno-Scientific Resilience: A National Strategy for Israel*, Special Report to the Prime Minister, eds. Isaac Ben-Israel, Eviatar Matania, and Leehe Friedman (in Hebrew) (September 2020), 32.
[37] COE CAHAI, "Towards Regulation," 130.

1. Fairness: striving for substantial equality, prevention of biases and discrimination (in information, in the process, and in the product), and avoidance of widening socioeconomic and educational gaps.
2. Accountability: incorporates the principles of transparency (information about the process and related decision-making); Explainability: being able to explain on the level of individual users, as well as on a collective level if the system affects groups; Ethical and legal responsibility: determining the responsibilities for setting reasonable measures to prevent the risk of harm.
3. Protecting human rights: preventing harm to life; Privacy: preventing damage to privacy due to collecting, analyzing, and processing information; Autonomy: maintaining the individual's ability to make intelligent decisions; Civil and political rights: right to elect, freedom of speech, and freedom of religion.
4. Cyber and information security: maintaining the systems in working order, protecting the information, and preventing misuse by a malicious actor.
5. Safety: preventing danger to individuals and to society.
6. Maintaining a competitive market and rules of conduct that facilitate competition.

The Ethics Report[38] then proposes the following model, to match different regulatory approaches based on the risk level associated with a particular activity. Thus, for example, high-risk activities are better addressed by legislation and self-regulation ex ante, than by post hoc judicial intervention. At the other end, low-risk activities do not necessarily require dedicated legislation, and can be addressed through standards and self-regulation. This model, of course, is not meant to apply in a rigid fashion. Rather, it presents a framework that enables policy makers and regulators to gauge the appropriate means of regulating an activity, factoring in a multitude of variables. The Report notes that the question of "who regulates" is no less important: regulation by a central AI body enables the development of consistent policies; however, there is a risk of overregulation and chilling innovation if a regulation is adopted across the board. Conversely, regulation could be left to different sector-based bodies, which would allow for greater experimentation, at the expense of uniformity of rules.[39]

1.4.2 Impact of AI on Human Rights

The Council of Europe and Commissioner for Human Rights issued a Recommendation involving steps to protect human rights from AI.[40] The Recommendation notes that the threat of AI to human rights is the central concern going forward. It suggests that public authorities perform human rights impact assessments "prior to the acquisition and/or development of [an AI] system" and that assessment must determine "whether an AI system remains under meaningful human control throughout the AI system's lifecycle."[41] The Recommendation also urges that "AI actors take effective action to prevent and/or mitigate the harms posed by their AI systems."[42] Furthermore, AI systems must not be "complex to the degree it does not allow for

---

[38] K. Nahon, A. Ashkenazi, R. Gilad Bachrach, D. Ken-Dror Feldman, A. Keren and T. Shwartz Altshuler, "Working Group on Artificial Intelligence Ethics & Regulation Report," in *The National Initiative for Secured Intelligent Systems*, 172.

[39] COE CAHAI, "Towards Regulation," 139.

[40] Council of Europe (COE) and Commissioner for Human Rights (CHR), "Unboxing Artificial Intelligence: 10 Steps to Protect Human Rights" (Recommendation) (May 2019).

[41] COE and CHR, "Recommendation," 7.

[42] COE and CHR, "Recommendation," 9.

human review and scrutiny"[43] and independent oversight should be required at the "administrative, judicial, quasi-judicial and/or parliamentary levels."[44] The Recommendation also prohibits the use of "AI systems that discriminate or lead to discriminatory outcomes," which includes "transparency and accessibility of information of the training data used in the development of an AI system."[45] In the area of data protection and privacy, it states that the "use of facial recognition technology should be strictly regulated."[46] The most meaningful protection, which would largely mitigate the fears of an AI takeover, is that "AI systems must always remain under human control, even in circumstances where machine learning or similar techniques allow for the AI system to make decisions independently."[47] Finally, at a societal level, governments should promote AI literacy through "robust awareness raising, training, and education efforts" and developers and appliers of AI should be required to gain knowledge of human rights law.[48]

1.4.3 *Trustworthy AI and Regulation by Design*

The High-Level Expert Group on Artificial Intelligence guide for creating *Trustworthy AI* focuses on three general areas of responsible AI – lawful AI, such as conformity with the GDPR; ethical AI, which is especially important when hard law rules are inexistent; and robust AI. Robust AI is a relatively vague concept that requires AI to "perform in a safe, secure and reliable manner, and safeguards should be foreseen to prevent any unintended adverse impacts."[49] The idea of designing safeguards to prevent any unintended adverse impacts is vital in principle, but as a general statement it is specious and without meaningful content. Finally, the study lists seven requirements for trustworthy AI: (1) human agency and oversight, (2) technical robustness and safety, (3) privacy and data governance, (4) transparency, (5) diversity, nondiscrimination, and fairness, (6) environmental and societal well-being, and (7) accountability.[50]

The High-Level Expert Group on Artificial Intelligence, discussed above, recognizes the technical component in the development of trustworthy or responsible AI. One element is the development of "whitelist" rules (acceptable or required behaviors or states) that the system should always follow and "blacklist" rules (restrictions on behaviors or states that the system should never transgress). This would be given to AI developers a priori in order to design systems that do not violate the prohibitions and incorporate the required safeguards.[51] An important component of any regulation of AI will be ex ante. Instead of waiting for problems to surface, regulation by design attempts to prevent the problems from occurring in the first place. Early examples of this coopting of technology for regulatory purposes are privacy-by-design and security-by-design. For example, the requirements of the GDPR should be incorporated into the design of an AI system. Thus, law and ethical principles would be used to standardize AI development.

---

43 COE and CHR, "Recommendation," 10.
44 COE and CHR, "Recommendation," 10.
45 COE and CHR, "Recommendation," 11.
46 COE and CHR, "Recommendation," 13.
47 COE and CHR, "Recommendation," 13–14.
48 COE and CHR, "Recommendation," 14.
49 European Commission's High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI" (2019), 6–7.
50 European Commission, "Trustworthy AI," 2.
51 European Commission, "Trustworthy AI," 14 and 21.

The US Government has advanced a strategic AI development plan.[52] It is characterized as a three-level structure. At the bottom are foundational values or principles that cut across all areas of AI innovation and applications, and include: (1) ethical, legal, and societal implications; (2) safety and security; (3) standards and benchmarks; (4) datasets and environments; and (5) capable AI workforce. Based on these foundational values, basic research and development focuses on two general areas: long-term investments and human–AI collaborations. Under the former category research is focused on data analytics, perception, theoretical limitations, general AI, scalable AI, human-like AI, robotics, and hardware. In the area of human–AI collaboration, the focus is targeted at human-aware AI, human augmentation, natural language processing, interfaces, and visualizations. Finally, in the application of AI, the plan recognizes the fields or sectors of agriculture, communications, education, finance, government services, law, logistics, manufacturing, marketing, medicine, science and engineering, transportation, and security. This strategy is well thought out but implementation of it will be problematic since it is based on collaboration across public-private entities, industries, and academic disciplines.

Regulation by design would not only consist of implementing existing law and ethical principles but would include a development process that anticipates future problems that may have negative ethical and socioeconomic impact. A component of regulation by design includes the allocation of responsibility. Currently, responsibility and potential liability can be easily allocated to the humans who develop and apply AI systems. This is because AI systems today are "closer to intelligent tools than sentient artificial beings."[53] However, "should the current predictions of superintelligence become realistic prospects, human responsibility alone" may not be sufficient. Instead, interdisciplinary assessments will be needed to determine where moral and legal responsibility lies when "AI participates in human-machine networks."[54]

1.4.4 Glance into the Future

The current and short-term progeny of AI in commerce and government has reduced costs to businesses and consumers and has effectuated more egalitarian benefits such as greater access to justice. Thus, the fear of the dangers attributed to AI have been inflated. There remains a large chasm between today's AI and that in the foreseeable future and the creation of artificial general intelligence – "a notional future AI system that exhibits apparently intelligent behavior at least as advanced as a person across the full range of cognitive tasks."[55] Even though superintelligence may be decades away, it is important that humankind begin forming an "environment (human) protection" impact study on how best to ensure that superintelligence works to enhance human existence and preserve human dignity.

The above notion of due diligence begins with the current understanding of AI and its applications. Today's regulators need to use the new technologies of today – machine learning, autonomous vehicles and systems, and AI decision-making – to develop frameworks and human capital necessary to deal with the AI of the future. Such frameworks will need to account for numerous factors, such as the quality and costs of certain technologies, as well as security, weaponization, privacy, safety and control, workforce, and fairness and justice concerns. The future regulator will be a technologist knowledgeable of the workings of AI, as well as being

---

[52] National Science and Technology Council, "AI Strategic Plan," 16.
[53] HUB4NGI, "Responsible AI."
[54] HUB4NGI, "Responsible AI."
[55] National Science and Technology Council, "Preparing for the Future," 7.

steeped in the understanding of the primary importance of democratic institutions and human dignity.

1.5 SCOPE OF COVERAGE

This section describes the book's coverage of a broad selection of topical areas with their own unique issues and problems. It provides a truly interdisciplinary and global perspective of the law and ethics of AI. The author group is a cosmopolitan mix of legal scholars, legal practitioners, and technologists in a variety of countries including Austria, China, Estonia, France, Germany, Italy, Japan, Netherlands, Spain, Switzerland, Turkey, United Kingdom, and United States.

The book is unique due to its breadth of coverage. It is divided into seven parts: Development and Trends; Contracting and Corporate Law; AI and Liability; AI and Physical Manifestations; AI and Intellectual Property Law; Ethical Framework for AI; and Future of AI. Part I introduces the key elements of AI and lays the foundation for the understanding of subsequent chapters. It includes an examination of the potential of AI to make law more efficient and less biased. It also examines the dangers of AI relating to its regulation, liability of entities that use AI, the replication of bias, and threats to democratic institutions. Chapter 2 is written by law and political scientists, as well as technologists who explain the various types of AI from machine learning to AI decision-making. Finally, the impact of AI and technology on the practice of law will be explored.

Part II consists of a series of chapters covering the application of AI to contracting and company law. In the area of contracting, the impact of AI on the negotiation, drafting, and formation of contracts, as well as in the performance of contracts, will be discussed. The final chapter of the part examines the role of AI in corporate decision-making and the board directors' duty of disclosure to shareholders.

Part III examines the issues of liability related to the creation and implementation of AI, including: a comparative analysis of the application of existing tort theories and potential liabilities from the European and American perspectives; an analysis of the question of liability relating to AI decision-making, data protection, and privacy; and an analysis of the application of agency law to AI systems.

Part IV focuses on the physical manifestations of AI, such as self-driving cars, other types of autonomous systems including robots, and the interconnectivity of AI to the Internet of Things. The scholars ask what happens if there are algorithmic errors that cause harm and who is liable for damages? The conclusion is that a new liability regime will be needed to allocate liability between the creators of the AI-controlled manifestations and those who sell or implement the AI system.

Part V examines the intersection of AI and intellectual property law. Key issues to be discussed include the patentability of AI from European and American perspectives; whether AI should be recognized as the creator of intellectual property; and whether AI-generated artistic works should be recognized under copyright law.

Part VI distinguishes between the ethical and unethical uses of AI. Given that regulation often lags behind technological developments, ethics will play an important part in setting limits for AI applications. The focus is on the relationship of AI to consumers in the areas of data privacy and security and the implications of AI for consumer law in general. The topics analyzed include whether AI should be recognized under the law as an artificial being, much like corporations; that is, should advanced AI be given legal status? Also studied are the implications of AI for legal and judicial ethics. How do current ethical standards apply to the lawyer's use of AI? The final

chapter theorizes that the best approach is moving beyond traditional approaches to ethics to a model of standardizing ethical AI.

The final part, Part VII, anticipates the future of AI as a disruptive force in such areas as the role of AI in the judicial system, public policy, legality and regulation of AI, and the ability of competition law to prevent AI collusion. The penultimate chapter ("The Folly of Regulating against AI's Existential Threat") ponders the future of AI. It takes a costs-benefits approach to the potential existential danger of advanced AI and suggests the appropriate government policy toward this accelerating technology. The final chapter summarizes the major findings and recommendations of the book.

Some of the more specific perspectives captured in the analysis include small and large business, government officials and regulators, legal practitioners and educators, ethicists, consumers, and citizens. Cross-chapter analyses cover the use of AI in government decision-making; legal practice (negotiation, drafting, and performance of contracts, as well as company law); ethical use of AI; and legal liability for AI including in tort law, data protection and privacy, and in agency law. Part VII also examines the issue of the liability for AI decision-making, liability for physical manifestations of AI, such as self-driving cars, other autonomous systems, robotics, and harm related to interconnectivity. The symbiotic relationship between AI and intellectual property law is explored including AI as inventor, patentability of AI, and protection of AI-generated works under copyright law.

From a broader perspective, there is the issue of "just because something can be done or achieved does that mean it should be done?" The normative element of autonomous systems and advanced AI are discussed in relationship to consumers, ethical frameworks, AI as a legal person, and control of AI through standardization. In Chapter 27, Professor John O. McGinnis leaves the reader with a positive and hopeful view of the development of advanced AI. He notes that AI as an existential threat to democracy and humanity is mostly speculative, but not certain. In the end, rationality warrants the encouragement of AI research because of the benefits it holds for humankind. AI is not the case for the application of a precautionary principle to prevent unexpected harm. It is simply a product of human creativity that can be harnessed for the greater good. Governments through funding and facilitative regulations or standardization should play a key role in this harnessing.

Essence of AI

What Is AI?

Pascal D. König, Tobias D. Krafft, Wolfgang Schulz, and Katharina A. Zweig

2.1 INTRODUCTION

As artificial intelligence (AI) systems increasingly become more widespread in society, many people regularly interact with them as a normal part of their everyday lives. The increasing prevalence of and seeming familiarity with everyday applications that employ AI may, however, easily betray its complexity. Diagnoses of the role of AI in society in public and scholarly discourse regularly depict AI as a uniform, monolithic phenomenon – almost like a force of nature that drives societal change. This is palpable, for instance, in statements that posit that AI will transform all aspects of social and economic life. However, while it is true that a certain set of technological advances largely rooted in computer science are responsible for an entire array of innovations in various domains, speaking of AI as a single technical entity conceals how elusive and multifaceted the term is.

Not only is there no commonly accepted definition of AI, but what the term refers to will, at the very least, depend on whether one is talking about AI as (a) a scientific field, (b) a technology or method, or (c) concrete applications of AI systems. The description of AI below will refer to each of these aspects and will partly show a certain overlap that is hard to avoid when talking about AI. We will use this categorization in the text for greater clarity.

A closer look at concrete implementations of AI quickly reveals a considerable diversity in its technical features, purposes, and scope. This means that the term AI is used differently by different people and may refer to many different things. Obviously, when it comes to confronting ethical, legal, and regulatory questions that are tied to the use of AI systems and their consequences, it is important to have a clear understanding of what one is dealing with. Adequate regulatory responses in particular may be hindered by a disconnect between policy actors and AI researchers in how they understand AI.[1] From a regulatory and policy perspective, it is therefore important to have an understanding of AI that captures aspects relevant to societal intervention while being aware of AI's current and future technical capabilities. Yet it is equally important to recognize that any attempt to get to the essence of AI needs to acknowledge that there is no common understanding of its definition and that it is used for multiple interwoven but separate concepts by different people. The present chapter thus aims to offer an answer to the question "what is AI?" that takes into account these considerations. It describes what makes

---

[1] P. M. Krafft, M. Young, M. Katell, K. Huang, and G. Bugingo, "Defining AI in policy versus practice," *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (New York: ACM, 2020), pp. 72–78.