

# Artificial Intelligence and the Law

## Editors

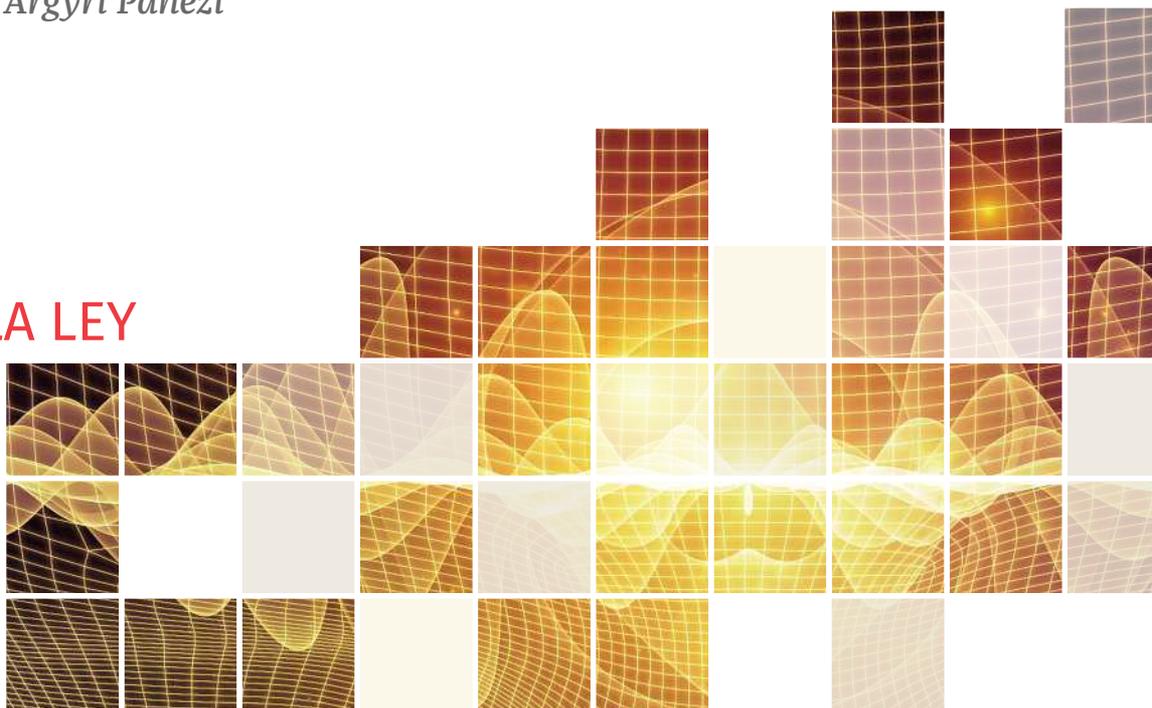
*Pablo García Mexía*  
*Francisco Pérez Bes*

## Coordinator

*Argyri Panezi*



■ LA LEY





# Artificial Intelligence and the Law

**Editors**

*Pablo García Mexía*

*Francisco Pérez Bes*

**Coordinator**

*Argyri Panezi*

© Authors, 2021

© Wolters Kluwer Legal & Regulatory Spain, S.A.

**Wolters Kluwer Legal & Regulatory Spain, S.A.**

C/ Collado Mediano, 9

28231 Las Rozas (Madrid)

**Tel:** 902 250 500 – Fax: 902 250 502

**e-mail:** clientes@wolterskluwer.com

<http://www.wolterskluwer.com>

**Edition:** June 2021

**Legal Deposit:** M-20026-2021

**ISBN print version:** 978-84-18662-40-9

**ISBN electronic version:** 978-84-18662-41-6

Design, Prepress and Printing: Wolters Kluwer Legal & Regulatory Spain, S.A.

*Printed in Spain*

© **Wolters Kluwer Legal & Regulatory Spain, S.A.** All rights reserved. Under the terms of art. 32 of Royal Legislative Decree 1/1996 of 12 April 1996, approving the Law on Intellectual Property, Wolters Kluwer Legal & Regulatory Spain, S.A. expressly objects any use of the contents of this publication without prior express authorization from the publishers, which includes in particular any reproduction, modification, recording, copying, exploitation, distribution, communication, transmission, sending, reuse, publication, processing or any other total or partial use in any form or by any means or format of this publication.

Any form of reproduction, distribution, public communication or transformation of this work can only be carried out with the express authorization of its owners, except for the exceptions provided by the Law. Please contact to **Cedro** (Centro Español de Derechos Reprográficos, [www.cedro.org](http://www.cedro.org)) if you need to photocopy or scan any part of this work.

The editor and the authors will not accept responsibility for the possible consequences caused to natural or legal persons who act or cease to act as a result of any information contained in this publication.

WOLTERS KLUWER LEGAL & REGULATORY SPAIN shall not be liable for the opinions expressed by the authors of the contents, as well as in forums, chats, or any other participation tools. the authors of the contents, as well as in forums, chats, or any other participation tools. Likewise, WOLTERS KLUWER LEGAL & REGULATORY SPAIN shall not be held responsible for any possible infringement of intellectual property rights that may be attributable to these authors.

WOLTERS KLUWER LEGAL & REGULATORY SPAIN is exempt from any liability for damages of any kind that may be caused by such authors. damages of any nature that may be due to the lack of truthfulness, accuracy, completeness and/or timeliness of the contents transmitted, disseminated, disseminated and/or updated of the contents transmitted, disseminated, stored, made available or received, obtained or accessed through its PRODUCTS. nor for the Contents provided or offered by third parties or entities. persons or entities.

WOLTERS KLUWER LEGAL & REGULATORY SPAIN reserves the right to eliminate any content which is content which is untrue, inaccurate and contrary to the law, morality, public order and good customs.

**Publisher's Note:** The text of the judicial decisions contained in the publications and products of **Wolters Kluwer Legal & Regulatory Spain, S.A.** is supplied by the Judicial Documentation Center of the General Council of the Judiciary (Cendoj), except for those that have been provided to us from time to time by the communications offices of the collegiate judicial bodies. The Cendoj is the only body legally empowered to collect such decisions. The processing of the personal data contained in these resolutions is carried out directly by the aforementioned body, since July 2003, with its own criteria in compliance with the regulations in force on the matter, being therefore its sole responsibility for any error or incident in this matter.

### 3. REACTIONS FROM THE LEGAL FIELD

Even the most sceptical voices with regard to the regulation of AI not only think that it is possible to act on it, but actually demand that it be done so<sup>(66)</sup>.

In fact, technological solutions are already beginning to abound, in the form of mathematical mechanisms or industry applications<sup>(67)</sup>. And also organizational formulas such as the identification of problematic contexts, the introduction of anti-bias tests or the promotion of investment in these areas<sup>(68)</sup>.

However, the experience of an intense relationship between digital technologies and the Law since the mid-nineties has taught an important lesson<sup>(69)</sup>. It is that, like any field of human action, legal rules will have to go into regulating certain aspects of AI as well.

And this regulation will take the form of either *soft-law*, or self-regulation by the industry itself, through ISO and similar standards and norms<sup>(70)</sup> or codes of good practice<sup>(71)</sup>; or “traditional” or “hard” legal norms, that is, State-supported, since, as we have seen, the principles and rights at stake (dignity, equality, privacy...) are, after all, of the utmost importance<sup>(72)</sup>.

If, in view of the already profuse literature on the legal regulation of AI, we were to highlight two principles above all others that have been proposed, there is no doubt that they would be the principle of *human centrality*, in the form of

(66) By all means, it is enough to mention the case of the American Center for Data Innovation, several times cited in this work.

(67) WHITTAKER, M. et al., *AI Now Report 2018*, p. 24-27, [https://ainowinstitute.org/AI\\_Now\\_2018\\_Report.pdf](https://ainowinstitute.org/AI_Now_2018_Report.pdf)

(68) MCKINSEY GLOBAL INSTITUTE, *Notes from the AI frontier: Tackling bias in AI (and in humans)*, 2019, p. 6, <https://www.mckinsey.com/~media/McKinsey/Featured%20Insights/Artificial%20Intelligence/Tackling%20bias%20in%20artificial%20intelligence%20and%20in%20humans/MGI-Tackling-bias-in-AI-June-2019.ashx>

Both types of measures can also be found in ICO, cit.

(69) CALO, R. underlines this necessary reflection in “Robotics and the Lessons of Cyberlaw”, *SSRN*, 2015, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2402972](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2402972)

(70) BERTOLINI, A. and PALMERINI, E., “Regulating Robotics. A Challenge for Europe”, EU Parliament, Directorate General for Internal Policies, *Workshop on Upcoming Issues of EU Law*, 2014, p. 194 ff. <http://www.robotlaw.eu/>

(71) *Algorithmic Accountability...*, cit., p. 24-25.

(72) *Algorithmic Accountability...*, cit., p. 24-25.

GUIHOT, M. et al., cit. suggest actions along the same lines, although, especially with respect to the so-called Big Techs, and in view of the flagrant imbalance of information that these accumulate in their favour, they uphold solutions based more on mere suggestion (*nudging*) than on openly binding measures.

control over intelligent systems (as, for example, Japan's Human Centric AI<sup>(73)</sup> initiative or the Council of Europe<sup>(74)</sup> advocate); and the principle of *accountability* of these systems, as advocated by the European Parliament among many other sources<sup>(75)</sup>. It must be pointed out that these two principles, like the others hereinafter referred to, are also included in the ethical declarations mentioned earlier, which does not detract in any way from their virtual parallel in the legal realm.

Human-centric AI and accountability are not values located on the same hierarchical plane. Human centrality (and the consequent control by AI) operates as a kind of supra-principle, insofar as the accountability of intelligent systems must be oriented precisely in the interests of such centrality and control. Accountability that is subordinate to the abuse of State power or exclusively for the benefit of private business would be useless. Moreover, it is this particular interlocking of human-centric AI and accountability that helps to generate another fundamental value in this area, namely *trust*: how can we expect people to trust intelligent systems that can act *legibus solutus*, regardless of any justification or consequences for the decisions they make?

What, by the way, should we understand by "accountability" of an intelligent system? The same source in the European Parliament specifies it in "the obligation to justify its decisions and the possibility of facing sanctions if such justification proves inadequate." This is a concept that is very close, in fact, to the classic obligation to justify their actions, which administrative law imposes on public authorities, both under Common and Civil law systems<sup>(76)</sup>.

---

(73) GOVERNMENT OF JAPAN, *Social Principles of Human Centric AI*, <https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf>

(74) Council of Europe, *Addressing the Impacts of Algorithms on Human Rights*, Draft Recommendation of the Committee of Ministers to member States on the human rights impacts of algorithmic systems, 2018, <https://rm.coe.int/draft-recommendation-of-the-committee-of-ministers-to-states-on-the-hu/168095eecf>

(75) EUROPEAN PARLIAMENT, *Understanding algorithmic...*, cit. Among the above-mentioned sources: DOSHI-VELEZ, F., KORTZ, M., et al., "Accountability of AI Under the Law: The Role of Explanation", 2019, <https://arxiv.org/abs/1711.01134>; R. Caplan et al., *Algorithmic Accountability...*, cit.; M. Whittaker et al., *AI Now Report 2018*, cit.; o Center for Data Innovation (CDI), *How Policymakers...*, cit.

(76) In the same vein, the definitions of the World Wide Web Foundation or the research centre Data & Society should be mentioned, although these two entities do so in this sense equivalent to "liability", as a need to repair damage. See World Wide Web Foundation, "Algorithmic Accountability, 2017, [http://webfoundation.org/docs/2017/07/Algorithms\\_Report\\_WF.pdf](http://webfoundation.org/docs/2017/07/Algorithms_Report_WF.pdf); CAPLAN, R. et al., *Algorithmic Accountability*, p. 22.

The principle of accountability has achieved primacy over three other adjacent principles, now absolutely topical in the already profuse literature on our subject:

— One is algorithmic *fairness*, which the best definitions equate to “absence of unwanted bias”<sup>(77)</sup>.

— The second is the principle of *transparency*, which the European Parliament again considers to be “availability of the system code and the relevant documentation”<sup>(78)</sup>.

— And the third, the so-called principle of *explainability*. Hand in hand with, among others, the Berkman Klein Center at Harvard University, we will understand it as “interpretability or comprehensibility of an intelligent system”<sup>(79)</sup>.

This last principle, explainability, is undoubtedly the most important of the three, insofar as it is worthy of enormous attention from legal writing and completely genuine of AI, in view of the well-known algorithmic opacity.

Among the very abundant literature on this principle, two ideas seem to stand out. One is that, although the algorithmic decision-making process is not human and is highly opaque, it does not seem reasonable to impose higher requirements on it than those generally provided for in today’s legal systems to regulate any other decision-making process<sup>(80)</sup>.

(77) EUROPEAN PARLIAMENT, *Understanding algorithmic...*, cit.

(78) EUROPEAN PARLIAMENT, *Understanding algorithmic...*, cit.

(79) DOSHI-VELEZ, F., KORTZ, M., et al., “Accountability of AI”, cit., p. 2 y 3. See also A. Campolo et al., *AI Now Report 2017*, p. 26, [https://ainowinstitute.org/AI\\_Now\\_2017\\_Report.pdf](https://ainowinstitute.org/AI_Now_2017_Report.pdf); and European Parliament, *Understanding algorithmic...*, cit., p. 30.

Three main techniques can be followed to implement the principle of explainability: a) The ‘black-box’ approach, which analyses the behaviour of the system so to speak without ‘opening the lid’, i.e. without any knowledge of its code. Explanations are built from observations of the relationships between system inputs and outputs. b) The ‘white-box’ approach: unlike the black-box approach, this one assumes that analysis of the system code is possible. An example of early work in this direction is the Elvira system for graphical explanation of Bayesian networks. c) And the constructive approach: in contrast to the first two, which assume that the system already exists, this approach consists of designing the intelligent system taking into account the requirements of explainability (“explainability by design”). See in this respect European Parliament, *Understanding algorithmic*, cit. p. IV.

(80) This is noted in an influential study by the European Commission, *Algo: aware Raising Awareness on Algorithms*, 2018, p. 23, <https://actuary.eu/wp-content/uploads/2019/02/AlgoAware-State-of-the-Art-Report.pdf>; and the Berkman Klein Center at Harvard University, in DOSHI-VELEZ, F., KORTZ, M., et al., “Accountability of AI”, cit.

The second idea is that we are not faced with a blind and mechanical possibility of demanding an explanation under all circumstances. As the Berkman Klein Center of Harvard itself points out, such an<sup>(81)</sup> explanation will only make sense if: a third party has suffered a damage that, in addition, is redressable (such redressable damage would, for example, be absent in the algorithm of a price comparator); or if there is some interest in the explanation in the face of well-founded suspicions of error in the system (due to inadequate inputs, unexplained results or distrust of the right intentions of the system).

Be that as it may, the principle of explainability comes with serious drawbacks. Thus some of the problems detailed by the aforementioned Center for Data Innovation are particularly important: Firstly, in view of the enormous complexity of some systems, even highly trained technologists may be unable to understand anything after accessing them. Second, some of these algorithms could be protected by copyright, given their status as software, which could legitimately lead their owners to refuse to provide any explanation. Thirdly, such explainability could end up benefiting malicious agents who access it. And fourth, and as a general rule, interpretability and accuracy are inversely related in AI, so reinforcing the former could undermine the immensely innovative potential of these technologies<sup>(82)</sup>. To the above-mentioned drawbacks, the highly authoritative American experts Annany & Crawford add some others, of which one will be highlighted, that is what they call “resilient transparency”, or the deliberate intention to camouflage important data with an additional flood of useless information<sup>(83)</sup>.

Because, in contrast to the principle of explainability, the principle of accountability operates, so to speak, “at a certain distance” from the technological code, it can circumvent each and every one of the disadvantages exposed and which are precisely related to the code. Furthermore, it overcomes the clear limitations of the fairness principle, which are unavoidable if we bear in mind that, as we have seen, all intelligent systems operate by definition with biases. On the other hand, it is general to understand transparency as a subsidiary principle, as an instrumental one, of accountability itself: the former would serve to strengthen the latter<sup>(84)</sup>. While the pair of forces “justification-sanction” allows us to effectively confront our key need: to make the person the centre of AI, subjecting it to our control. This is what explains the primacy of accountability

---

(81) DOSHI-VELEZ, F., KORTZ, M., et al., “Accountability of AI”, cit.

(82) CENTER FOR DATA INNOVATION (CDI), *How Policymakers...*, cit., p. 9-13.

(83) ANANNY, M. and CRAWFORD, K., “Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability”, 2016, *New Media & Society*, 1-17, p. 7.

(84) DOSHI-VELEZ, F., KORTZ, M., et al., “Accountability of AI”, cit.; y ANANNY, M. and CRAWFORD, K., “Seeing without...”, cit., p. 2.

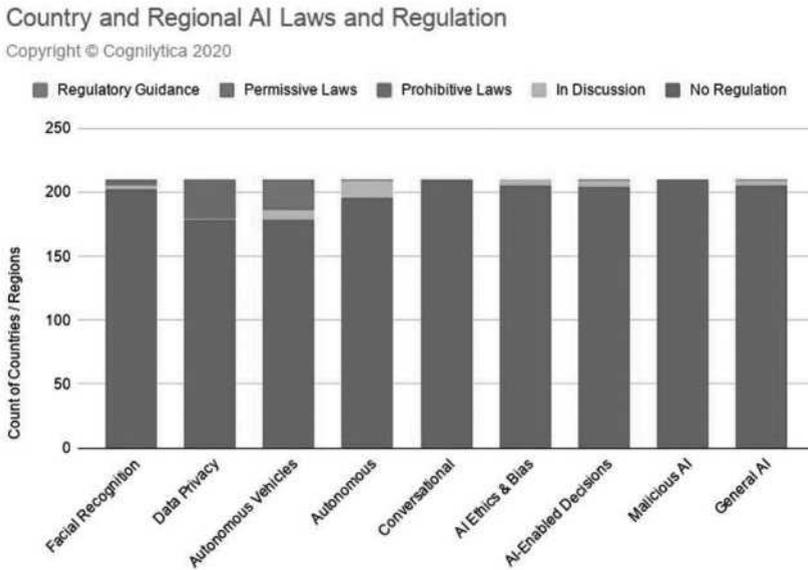
over its adjacent principles: fairness, transparency and even the “very attractive” one of explainability<sup>(85)</sup>.

All these principles ought to end up being part of legal provisions. And, of course, on the basis of such provisions, of case law. Until now, however, and with the sole exception of autonomous vehicles (as the following figure shows), this has only begun to happen in one specific area, that of privacy<sup>(86)</sup>.

(85) It is therefore not surprising that the very few countries that have begun to regulate AI, or to consider doing so, focus their initiatives precisely on the principle of accountability. This is the case with two legislative proposals of identical text and title, the *Algorithmic Accountability Act*, which in April 2019 were introduced in the US Senate and House of Representatives <https://www.congress.gov/bill/116th-congress/house-bill/2231/all-info>.

Instead of focusing on accountability, the *Executive Order on Maintaining American Leadership in Artificial Intelligence* of 11 February 2019 pivots instead on the principle of trust, as a guarantee for innovation in AI that at the same time respects “civil rights, privacy and American values” (section 6). See <https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/>. This regulatory framework is the basis for the 10 principles on AI developed by the US White House Office of Science and Technology Policy, published in early January 2020 and also governed by the principle of “public trust”; these principles should guide the regulation that executive branch agencies subsequently seek to impose on the private sector. See <https://www.technologyreview.com/2020/01/07/130997/ai-regulatory-principles-us-white-house-american-ai-initiative/>. This is also the case with the European Commission’s Communication of 8 April 2019, COM(2019) 168 final, *Artificial Intelligence for Europe*, which calls for progress in this field with full respect for “the EU’s values and fundamental rights” (but also insists on the need for a “human-centric AI”). Also the Report of the High Level Expert Group on Artificial Intelligence appointed by the European Commission itself, of 8 April 2019, calls for the construction of a “trusted” artificial intelligence. In the same vein, the Declaration of the (then) 28 EU States on Cooperation in Artificial Intelligence, signed between April and July 2018 (although it mentions at the same level the principle of transparency), and the Italian *Libro Bianco sull’Intelligenza artificiale...*, cit. On all this, see <https://www.loc.gov/law/help/artificial-intelligence/europe-asia.php>

(86) A pioneering Dutch ruling from the Rechtbank Den Haag (The Hague, NL), dated 5 February 2020, is worth mentioning in this regard. This ruling states that the intelligent algorithm of the so-called ‘SyRI’ system, used by the Dutch authorities to predict which people would be most at risk of committing public housing or social security fraud, is contrary to the right to privacy (Article 8(2) of the EU Charter of Fundamental Rights, CFREU). The algorithm compiled the profiles of past offenders in order to achieve “offender patterns”. From these patterns, the system tracked databases to identify which individuals were most likely to fit into these predetermined patterns and thus be more closely monitored. The Court considers all this to be contrary to the right to privacy, since such monitoring would not be motivated by any other reason than the stigmatization derived from the SyRI algorithm. This is the first time that a European court has outlawed the operation of an intelligent algorithm in the light of European fundamental rights law. However, it is worth asking whether, even if the Court does not mention this other issue at all, SyRI would not also be contrary to the principle of equality, which is of course also provided for in the CFREU, in that it would have discriminated against citizens who were specially monitored on the sole basis of this algorithmic indication, which does not seem to be a sufficiently objective or reasonable justification, especially since those citizens were



**Figure 2.** The global regulatory landscape on AI.

Source: Worldwide AI Laws and Regulations 2020.

<https://www.cognilytica.com/2020/02/14/worldwide-ai-laws-and-regulations-2020/>

All of which leads us to briefly comment on our leading privacy rule, the EU’s GDPR. In addition to the one just mentioned, there are two reasons for this. The first is that the GDPR is the first legal norm in the world to regulate the impact of AI on privacy<sup>(87)</sup>. The second reason is that the GDPR can be considered the regulatory paradigm of the two key principles that we have been analysing, that of centrality or control by the person and that of accountability, both being at the same time the two basic pillars on which this rule is based.

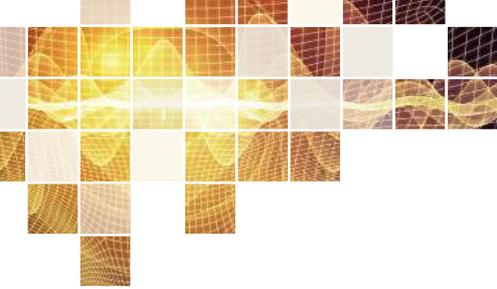
The principle of human centrality or control over intelligent systems is implemented in the GDPR through two channels. One, the right granted to data subjects not to be subject to decisions based solely on automated processing, including profiling, which significantly affect them, with or without legal effects<sup>(88)</sup>. In other words: a “right to minimum human intervention” in decisions with an impact on privacy.

generally found to be economically disadvantaged. For the ruling in Dutch, see <https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:865&showbutton=true>; for a good journalistic summary in English, see <https://www.dutchnews.nl/news/2020/02/governments-fraud-algorithm-syri-breaks-human-rights-privacy-law/>

(87) Countries as relevant as Canada are also beginning to follow these regulatory paths: see [https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/consultation-ai/pos\\_ai\\_202001/](https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/consultation-ai/pos_ai_202001/)

(88) Art. 22 GDPR.





It is now commonplace to say that artificial intelligence (AI) makes up one of the key technological advances for humanity. And yet, it is also general to state that AI represents one of the world's critical challenges from social, ethical and legal perspectives.

It is precisely in these mixed views that the book finds its deepest *raison d'être*. If societies are at least significantly sceptical about the impact of AI, it seems more than justified that those components of AI that may in principle put more social risks should find an ethical and legal counterweight. Indeed, this book aims to contribute to enrich the debate about the ethical and especially legal determinants of AI technologies.

The work is characterised by three main notes: it is multidisciplinary, international and introductory. Multidisciplinary, because it focuses mainly on legal aspects, but it does not disregard the technological one, obviously essential in view of its subject-matter; although its readers may not be technologists, it is decisive for them to be provided with the foundations for any feasible ethical, business or legal analysis. It is international, as its authors come from universities and other entities on three continents (the Americas, Asia and Europe), which has helped to approach the issues from different territorial, thus complementary perspectives. Thirdly, it is an introductory book to the main ethical and legal (also business) problems of AI; at recording all of them, though, it will provide the reader with the general vision often sought when first approaching a matter or when trying to identify its very essence.

