Consumer protection requires artificial intelligence

Lippi, M., Contissa, G., Lagioia, F. *et al.* Consumer protection requires artificial intelligence. *Nature Machine Intelligence* 1, 168–169 (2019).
Published 25 March 2019

https://doi.org/10.1038/s42256-019-0042-3

Consumer Protection Requires Artificial Intelligence

Marco Lippi[1,*], Giuseppe Contissa[2], Francesca Lagioia[2,3], Hans-Wolfgang Micklitz[3], Przemysław Pałka[4], Giovanni Sartor[2,3], and Paolo Torroni[5]

[1]DISMI, University of Modena and Reggio Emilia
[2]CIRSFID, University of Bologna
[3]Law Department, European University Institute
[4]Center for Private Law, Information Society Project, Yale Law School
[5]DISI, University of Bologna
[*]Corresponding author: Marco Lippi, marco.lippi@unimore.it

Modern algorithmic technologies for large-scale analysis of user-related data enable businesses to profile their customers and to anticipate and influence their behaviour, threatening, by doing so, their consumer rights [1]. We argue that in the consumer domain, regulatory initiatives [2] are necessary, but insufficient to ensure effective consumer protection. Legal interventions must be complemented by the countervailing power of civil society, namely of consumer organizations and individual consumers. This idea was famously proposed by economist Ken Galbraith, who argued that "an opposing exercise of power is the principal solvent of economic power, the basic defense against its exercise in economic affairs" [3]. We endorse this idea, with an addition: in the era of Artificial Intelligence (AI), when the power of traders is boosted by AI, an effective countervailing power needs also to be supported by AI.

Bringing AI to the side of consumers

A few examples of consumer-empowering technologies are already with us, including ad-blocking systems, anti-spam software and anti-phishing techniques. Yet, there is a need to move forward. Like companies, consumers and their organizations too could benefit from the application of AI technologies to big data. On the side of consumers, services could be deployed with the goal of analyzing and summarizing massive amounts of product reviews [4], detecting discrimination in commercial practices [5], recognizing identity fraud [6], or comparing prices across a multitude of platforms.

AI could help fill the gap between law in the books and law in action, between the assertion of rights and their practice and enforcement. Consumer law, established to counter the power imbalance between traders and consumers, gives the latter a wide range of legal tools to resist abuse by traders. Those include, among others, regulations against unfair contract terms, and against unfair, misleading and aggressive commercial practices. In addition, data protection law, and in particular the recently adopted European General Data Protection Regulation (GDPR), gives consumers and their organizations the possibility to react against abuses involving the processing of consumer data. Unfortunately, consumers usually lack the factual capacity and the resources to make use of existing regulatory tools. AI-based technologies could contribute to address this shortcoming.

One particular area where progress could be made is consumer contracts. In fact, such contracts are almost never read by subscribers, who simply agree without knowing what they are accepting [7]. In [8] we presented the result of an analysis of fifty online contracts, showing that about one sentence in ten contains a potentially unlawful clause, including unilateral changes of service conditions, unfair arbitration clauses, etc. Machine learning and natural language processing technologies can help consumers to detect unlawful clauses, by analyzing contractual documents in order to validate their content. One example in this direction is offered by CLAUDETTE,[1] a web server for the automatic detection of potentially unfair clauses in online Terms of Service. More generally AI-based technologies for textual analysis may provide new powerful tools to consumers and their organizations. For example, such technologies have been applied to privacy policies, to assess compliance with GDPR [9], although privacy documents are far more complex than consumer contracts. The growing interest in this area resulted in several proposals for automatically extracting, categorizing, and summarizing information from privacy documents, and assisting users in processing and understanding their contents [10]. A major effort in this direction is being carried out within the Usable Privacy Policy project [11].[2]

Automatic tools for empowerment

The mentioned tools are a promising start. However, we still have a long way to go before a repertoire of AI applications that effectively empower civil society and consumers can be provided. At this stage a crucial step is the identification of relevant research areas and challenges ahead.

A first domain is the analysis of textual documents. Consumers usually do not read, and in any case are unable to negotiate the documents prepared by traders. Therefore, legal rules have been put in place to structure and constrain the content of various documents directed to consumers, such as terms of service, privacy policies, and other specific contracts (e.g. in the banking, transport, and holiday sectors). Systems capable to perform a preliminary analysis and legal

[1] `https://claudette.eui.eu/demo`
[2] `https://www.usableprivacy.org`

evaluation of these documents, assessing their compliance with the applicable legal rules, could provide valuable support to consumers and their organizations, by identifying unlawful clauses (or omissions) and by attempting an initial legal analysis.

A second domain is the detection of practices that may infringe consumer rights or negatively affect their interests. For instance, consumers are often tracked, profiled or subjected to automated decision without being fully informed on the processing of their data, or without being able to easily access this information. Similarly, consumers are often oblivious targets of unfair, aggressive or discriminatory commercial practices, as well as of frauds and security violations. Here, there is a need for advancing the research and technologies for the detection of identity frauds and anomalous situations, and for tracking and identifying unlawful uses of personal data (for example, see the recent discussion regarding online photos scraped without consent [12]).

An important related area is the detection and disabling of covert communication [13], and the reporting of potential information leaks, for example in mobile applications [14]. Moreover, because consumers alone would not be able to leverage the potential of such AI tools, a new class of enabling tools is needed, to facilitate the communication with NGOs and supervisory authorities, and possibly even class actions. A further area is thus the automated notification and large-scale processing of consumer-to-business data trails pointing to malpractices against consumer rights and interests. Many of the aforementioned consumer-empowering tools could become building blocks of "Privacy Digital Assistants", i.e., intelligent software agents acting on behalf of their owners, able to detect and notify privacy risks and violations, as well as to detect and disable privacy intrusive default settings and covert communication, by filtering information according to consumer preferences, and to report information leaks. Interesting proposals in this vein have been made by several groups [15, 16].

Lastly, the recent prospects opened by data-driven decision making have to some extent eclipsed the invaluable resource advocacy groups indisputably have: expert domain knowledge. We strongly believe that effective AI-based consumer-empowering tools should instead exploit such a knowledge. In particular, research efforts should focus more on the combination of knowledge-based and data-driven methods, for example by building on recent advances in neural symbolic and statistical relational learning, and by pushing for their applications in the legal domain.

An inter-disciplinary challenge

If AI tools supporting consumers and civil society are socially beneficial and technically feasible, we may wonder why they are a vision and not a reality. One of the main reasons is that building such tools requires both interdisciplinary cooperation and relevant investment.

Interdisciplinary cooperation should increase, first of all in academia. Even though the community of tech-savy lawyers and law-knowledgeable engineers

is growing,[3] there is not enough emphasis on consumer-focused research initiatives. An incentive for lawyers and engineers to cooperate could also consist in funding directed to AI & Law research, in particular in the domain of consumer empowerment. With regard to investment, we believe that currently private operators do not have sufficient incentive to develop consumer-friendly technologies. Even changes in legislation, such as the granting of punitive damages or a right to class actions, are unlikely to provide sufficient financial returns for this kind of investment.

Therefore, policy action is needed. Governments, foundations, or other stakeholders should facilitate research in these domains both by acknowledging the need, and by offering initial funding for legal-tech projects which aim at empowering consumers. In this way, there would be incentives and encouragement for interdisciplinary groups to form, both because scholars would realize that they can do it — and, in many ways, those are really fascinating projects — and because they could afford it. Finally, academics should communicate and work together with the practitioners and consumer organizations on the field. It is them who understand what precisely the problems are, who often have (access to) the valuable data, and who will, in the end, use these tools. Hence, the need to create applications that meet the actual requirements of consumer protection.

Conclusion

As AI is becoming truly ubiquitous, not only businesses but also consumers should enjoy its power. We envision a future where business endowed with "business-empowering" AI face consumers supported by "consumer-empowering" AI. These consumer-empowering technologies would contribute to reduce unfair and unlawful market behavior, and favor the development of legal and ethical business models. This would certainly be beneficial to individual consumers, and to their organizations, but it would also contribute to the better functioning of digital markets.

Acknowledgments

This work was financially supported by the European University Institute projects ARTSY (`artsy.eui.eu`) and CLAUDETTE (`claudette.eui.eu`).

Corresponding authors

Correspondence to Marco Lippi.

[3] See for instance the Future of Law conference, `https://conferences.law.stanford.edu/futurelaw/`

Competing Interests

The authors declare no competing interests.

References

[1] Shoshana Zuboff. *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs, 2019.

[2] Mireille Hildebrandt. Primitives of legal protection in the era of data-driven platforms. *Georgetwon Law Technology Review*, 2, 2018.

[3] John Kenneth Galbraith. *The Anatomy of Power*. Houghton-Mifflin, 1983.

[4] Nan Hu, Noi Sian Koh, and Srinivas K Reddy. Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales. *Decision support systems*, 57:42–53, 2014.

[5] Salvatore Ruggieri, Dino Pedreschi, and Franco Turini. Data mining for discrimination discovery. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 4(2):9, 2010.

[6] Shan-Hung Wu, Man-Ju Chou, Chun-Hsiung Tseng, Yuh-Jye Lee, and Kuan-Ta Chen. Detecting in situ identity fraud on social network services: A case study with facebook. *IEEE Systems Journal*, 11(4):2432–2443, 2017.

[7] Jonathan A Obar and Anne Oeldorf-Hirsch. The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society*, pages 1–20, 2018.

[8] Marco Lippi, Przemysław Pałka, Giuseppe Contissa, Francesca Lagioia, Hans-Wolfgang Micklitz, Giovanni Sartor, and Paolo Torroni. CLAUDETTE: an automated detector of potentially unfair clauses in online terms of service. *Artificial Intelligence and Law*, Feb 2019.

[9] Giuseppe Contissa, Koen Docter, Francesca Lagioia, Marco Lippi, Hans-W Micklitz, Przemysław Pałka, Giovanni Sartor, and Paolo Torroni. CLAUDETTE meets GDPR. Automating the evaluation of privacy policies using artificial intelligence. 2018. Study Report, Funded by The European Consumer Organisation (BEUC).

[10] Przemysław Pałka and Marco Lippi. Big data analytics, online terms of service and privacy policies. 2019.

[11] Shomir Wilson, Norman Sadeh, Noah A. Smith, Florian Schaub, Frederick Liu, Kanthashree Mysore Sathyendra, Daniel Smullen, Sebastian Zimmeck, Rohan Ramanath, Peter Story, and Fei Liu. Analyzing privacy policies at scale: From crowdsourcing to automated annotations. *ACM Transactions on the Web*, 13:1–29, 12 2018.

[12] O. Solon. https://www.nbcnews.com/tech/internet/facial-recognition-s-dirty-little-secret-millions-online-photos-scraped-n981921. *NBC News*, 12 March 2019.

[13] Julia Rubin, Michael I Gordon, Nguyen Nguyen, and Martin Rinard. Covert communication in mobile applications (t). In *Automated Software Engineering (ASE), 2015 30th IEEE/ACM International Conference on*, pages 647–657. IEEE, 2015.

[14] Michael I Gordon, Deokhwan Kim, Jeff H Perkins, Limei Gilham, Nguyen Nguyen, and Martin C Rinard. Information flow analysis of android applications in droidsafe. In *NDSS*, volume 15, page 110, 2015.

[15] Karuna P Joshi, Aditi Gupta, Sudip Mittal, Claudia Pearce, Anupam Joshi, Tim Finin, et al. Alda: Cognitive assistant for legal document analytics. In *AAAI Fall symposium*, volume 2016, 2016.

[16] Anupam Das, Martin Degeling, Daniel Smullen, and Norman M. Sadeh. Personalized privacy assistants for the internet of things: Providing users with notice and choice. *IEEE Pervasive Computing*, 17(3):35–46, 2018.