

Volontà e autonomia dell'Intelligenza Artificiale. Il nodo dell'imputazione della decisione algoritmica

BENEDETTA MARIA SABATINO

ABSTRACT

La ricerca è orientata a valutare la compatibilità dell'attuale sistema normativo e codicistico rispetto alle implicazioni della decisione assunta dall'intelligenza artificiale (specie allorché essa sia imprevedibile ed intellegibile) e ciò tanto nell'ambito del diritto contrattuale quanto nel sistema aquiliano.

Il lavoro approfondisce il tema della decisione algoritmica rispetto all'elemento volitivo su cui riposa il sistema contrattuale delineato dal Codificatore del 1942, nonché, laddove tale procedimento decisionale dell'I.A. sia svolto in campo extracontrattuale, in che modo il sistema della responsabilità sia suscettibile di applicazione dinanzi alle nuove sfide derivanti dall'impiego di simili tecnologie.

L'obiettivo, pertanto, è mettere in luce come la decisione autonoma dell'intelligenza artificiale, seppure nelle sue diverse "epifanie", possa essere osservata sia come manifestazione di volontà, quasi a far ritenere che l'I.A. possa in qualche modo essere qualificata come parte contraente nell'ambito dell'operazione contrattuale, sia come comportamento atto a provocare un danno, tanto all'utilizzatore quanto ai terzi. L'analisi, dunque, è rivolta alla menzionata componente volitiva dell'intelligenza artificiale ed ai conseguenti criteri di imputazione della decisione.

Il Capitolo I muove da alcune preliminari note introduttive relative alla storia dell'intelligenza artificiale ed agli sviluppi che si possono attendere nei prossimi anni, al *framework* normativo che sta sviluppandosi in ambito eurounitario e, con precipuo riferimento alla disciplina multilivello in materia di protezione dei dati personali (con particolare riguardo al Regolamento europeo 679/2016), alla tematica della discriminazione connessa alla decisione algoritmica e alla correlata problematica della c.d. *black box* tipica dei sistemi in esame, ponendo l'accento sul divieto di decisioni automatizzate sancito dall'art. 22 GDPR.

Il Capitolo II declina il problema della decisione dell'I.A. in ambito contrattuale: le considerazioni muovono invero dalla preliminare distinzione tra "*smart contracts*" ed "*algorithmic contracts*", per poi approfondire il rapporto con gli istituti dell'arbitraggio e della rappresentanza e sostenere, in ultima analisi, che colui che beneficia delle funzionalità delle nuove tecnologie debba considerarsi al tempo stesso onerato di sopportare i rischi derivanti dall'assunzione di una decisione non voluta, non potendosi trincerare dietro lo schermo dell'autonomia decisionale del *software*.

La questione dell'imputazione degli effetti del comportamento dell'I.A. emerge altresì nel Capitolo III, che si apre con l'analisi del tema della soggettivizzazione dell'intelligenza artificiale, che, invero, si pone in linea con le recenti tendenze dirette a riconoscere la personalità ad animali ed altre entità non umane.

In realtà, anche dall'esame dei provvedimenti recentemente assunti in sede eurounitaria, il modello di responsabilità per i danni cagionati dall'I.A. sembra essere fondato sul (ed orientato verso il) criterio del rischio, che caratterizza la società industriale ma – si può affermare – ancor di più quella digitale. Tuttavia, a differenza di quanto emerge dal *framework* europeo, in cui siffatto sistema di allocazione è destinato ai soli operatori, il lavoro, anche in ottica di una responsabilizzazione dell'utenza, tende a mettere in rilievo come il modello di responsabilità oggettiva appaia altresì declinabile per talune ipotesi dannose, imputabili a chi beneficia dell'attività di uno *smart product*, anche se nelle vesti di consumatore.

In definitiva, fintantoché non svilupperà una propria coscienza, l'intelligenza artificiale non è in grado di mettere in crisi le categorie tradizionali del diritto civile, atteso che l'attenzione dell'interprete deve essere rivolta non tanto al grado di intelligenza detenuto dal sistema di I.A., quanto alla strumentalità della sua attività rispetto ai bisogni dell'uomo.

The research is aimed at evaluating the compatibility of the current regulatory and codicistic system in relation to the implications of the decision taken by artificial intelligence (especially when it is unpredictable and intelligible) and this both in the context of contract law and in the Aquilian system.

The work examines the topic of the algorithmic decision with reference to the volitional element on which the contractual system outlined by the 1942 legislator rests, as well as, where this decision-making procedure of the A.I. is carried out in the non-contractual field, how the liability system is susceptible of application in the face of the new challenges arising from the use of similar technologies.

The goal, therefore, is to highlight how the autonomous decision of artificial intelligence, albeit in its various “epiphanies”, can be observed both as a manifestation of will, as if to suggest that A.I. may in some way be qualified as a contracting party in the context of the contractual operation, both as a behaviour capable of causing damage, both to the user and to third parties. So, the analysis is directed towards the aforementioned volitional component of artificial intelligence and the consequent criteria for imputing the decision.

Chapter I starts from some preliminary introductory notes about the history of artificial intelligence and the developments that can be expected in the coming years, to the regulatory framework that is developing in the EU area and, with particular reference to the multilevel legislation on protection of personal data (especially with regard to the European Regulation 679/2016), to the issue of discrimination connected to the algorithmic decision and to the related problem of the so-called black box typical of these systems, placing the emphasis on the prohibition of automated decisions enshrined by art. 22 GDPR.

Chapter II declines the problem of the A.I.’s decision in the contractual field: considerations actually start from the preliminary distinction between “smart contracts” and “algorithmic contracts”, to then deepen the relationship with the arbitration and agency institutions and ultimately argue that the one who benefits from the functionality of new technologies must be considered at the same time burdened to bear the risks arising from the taking of an unwanted decision, not being able to hide behind the screen of the decision-making autonomy of the software.

The issue of imputation of the effects of the A.I.’s behaviour also emerges in Chapter III, which opens with the analysis of the subjectivization of artificial intelligence, which, indeed, is in line with recent trends aimed at recognizing the personality of animals and other non-human entities.

In fact, even from the examination of the measures recently taken at the EU level, the responsibility model for the damage caused by the A.I. seems to be based on (and oriented towards) the risk criterion, which characterizes the industrial society but – it can be said – even more so the digital one. However, unlike what emerges from the European framework, in which such allocation system is intended only for operators, the work, also with a view to make users accountable, tends to highlight how the strict liability model also appears to be declinable for certain harmful hypothesis, attributable to those who benefit from the activity of a smart product, even if in the guise of a consumer.

Ultimately, as long as it does not develop a conscience of its own, artificial intelligence is not able to undermine the traditional categories of civil law, given that the interpreter’s attention must be directed not so much to the degree of intelligence held by the A.I. system, but to the instrumentality of its activity with respect to the man’s needs.