

La decisione nel prisma dell'intelligenza artificiale

a cura di
ERMANN0 CALZOLAIO

ERMANN0 CALZOLAIO (a cura di), *La decisione nel prisma dell'intelligenza artificiale*

LA DECISIONE NEL PRISMA DELL'INTELLIGENZA ARTIFICIALE

a cura di
ERMANN0 CALZOLAIO

La presente pubblicazione è finanziata dal Dipartimento di Giurisprudenza dell'Università di Macerata nell'ambito della ricerca "Decision and New Technologies" (DANT), che costituisce una delle iniziative di implementazione del progetto "Dipartimento di Eccellenza", finanziato dal MIUR, il cui tema generale è "Diritto e Innovazione: Europa e Cina di fronte alle sfide della globalizzazione"

Come da accordi con il Dipartimento di Giurisprudenza dell'Università di Macerata, questa copia è concessa dall'Editore per la pubblicazione *Open Access* nell'archivio dell'Università degli Studi di Macerata, nonché su altri archivi istituzionali e di ricerca scientifica ad accesso aperto.

Copyright 2020 Wolters Kluwer Italia S.r.l.
Via dei Missaglia n. 97 - Edificio B3 - 20142 Milano

I diritti di traduzione, di memorizzazione elettronica, di riproduzione e di adattamento totale o parziale, con qualsiasi mezzo (compresi i microfilm e le copie fotostatiche), sono riservati per tutti i Paesi.

Le fotocopie per uso personale del lettore possono essere effettuate nei limiti del 15% di ciascun volume/fascicolo di periodico dietro pagamento alla SIAE del compenso previsto dall'art. 68, commi 4 e 5, della legge 22 aprile 1941, n. 633. Le riproduzioni diverse da quelle sopra indicate (per uso non personale - cioè, a titolo esemplificativo, commerciale, economico o professionale - e/o oltre il limite del 15%) potranno avvenire solo a seguito di specifica autorizzazione rilasciata da EDISER Srl, società di servizi dell'Associazione Italiana Editori, attraverso il marchio CLEARedi Centro Licenze e Autorizzazioni Riproduzioni Editoriali. Informazioni: www.clearedi.org.

L'elaborazione dei testi, anche se curata con scrupolosa attenzione, non può comportare specifiche responsabilità per eventuali involontari errori o inesattezze.

INDICE

Autori.....	XI
-------------	----

INTELLIGENZA ARTIFICIALE ED AUTONOMIA DELLA DECISIONE: PROBLEMI E SFIDE

Ermanno Calzolaio

1. Spunti introduttivi.....	1
2. Le innovazioni tecnologiche e il diritto.....	3
3. La decisione quale criterio di osservazione del fenomeno.....	5

AI-BASED DECISION SUPPORT SYSTEM: FROM THEORETICAL BACKGROUND TO REAL WORLD APPLICATIONS

Emanuele Frontoni – Marina Paolanti

1. Introduction.....	9
2. Description of real-world applications.....	13
2.1. An Infrastructure for decision-making to support neonatal clinical care and research.....	13
2.2. Decision Support System for assisted reproductive technology.....	16
2.3. Visual and textual sentiment analysis of social media pictures using deep convolutional neural networks.....	17
2.4. Deep understanding of shopper behaviours and interactions using computer vision.....	19
3. Conclusion.....	21

CHI DECIDE? INTELLIGENZA ARTIFICIALE E TRASFORMAZIONI DEL SOGGETTO NELLA RIFLESSIONE FILOSOFICA

Carla Canullo

1. La filosofia moderna, la nascita del soggetto e il sogno della macchina.....	25
---	----

2.	<i>La contrée des philosophes</i>	27
3.	L'apparire inatteso della volontà.....	30
4.	La sfida della libertà.....	32
5.	La libertà della decisione.....	35

THE JUDICIAL DECISION BETWEEN LEGAL GAPS AND
TECHNOLOGICAL INNOVATION: SOME SUGGESTIONS FROM
THE 19TH AND 20TH CENTURIES

Monica Stronati

1.	Artificial intelligence between opportunities and challenges.....	37
2.	The boundaries of liability: from fault-based liability to strict liability and back (to fault-based liability)....	41
3.	Legal certainty, judicial predictability and Justice.....	51

LÉGIFÉRER SUR L'INCONNU.
À PROPOS DES SYSTEMES D'ARMES LÉTAUX AUTONOMES

Nicolas Haupais

1.	Introduction.....	57
2.	Que faire ? Comment légiférer ?.....	61
3.	Les aspects juridiques de l'interdiction/limitation de l'emploi.....	71
4.	La dynamique de l'interdiction: réciprocité et obliga- tions interdépendantes.....	74

LA DÉCISION DU CONTRACTANT AU PRISME DE
L'INTELLIGENCE ARTIFICIELLE

Garance Cattalano

1.	Introduction.....	77
2.	L'automatisation de la décision de conclure le con- trat.....	81
	A. Les modalités pratiques.....	81
	B. Les difficultés théoriques.....	82
3.	L'automatisation des décisions pendant l'exécution du contrat.....	85
	A. La simplification.....	85
	B. La rigidification.....	88

LA LÉGITIMITÉ DES DÉCISIONS CONTRACTUELLES
ÉMANANT D'UNE INTELLIGENCE ARTIFICIELLE

Frédéric Dournaux

1.	Introduction.....	91
2.	La prise de décision contractuelle par une intelligence artificielle.....	93
	A. L'intelligence artificielle, aide à la décision contractuelle.....	93
	B. L'intelligence artificielle, acteur de la décision contractuelle.....	95
3.	De quelques questions posées par la prise de décision contractuelle par une intelligence artificielle.....	96
	A. Les insuffisances de la technique.....	97
	B. Les risques relatifs à l'externalisation.....	98
	C. Les réserves d'ordre axiologique.....	99

DRIVERLESS CARS:
I NUOVI CONFINI DELLA RESPONSABILITÀ CIVILE
AUTOMOBILISTICA E PROSPETTIVE DI RIFORMA

Stefano Pollastrelli

1.	Premessa.....	103
2.	Dal conducente al robot. La categoria dei <i>Self Driving Vehicles</i>	107
3.	Il crepuscolo dell'art. 2054 c.c. Verso una nuova visione della responsabilità civile automobilistica.....	109
4.	Considerazioni conclusive.....	116

LA DÉCISION EN DROIT ET CONTENTIEUX
ADMINISTRATIFS FRANÇAIS

Pierre Serrand

1.	Introduction.....	119
2.	La décision en contentieux administratif: l'acte administratif décisoire.....	123
	A. La notion.....	123
	B. Les distinctions.....	126
3.	La décision en droit administratif: l'acte administratif discrétionnaire.....	130
	A. Décision et autorité administrative.....	131
	B. Décision et juridiction administrative.....	134

BIG DATA, DECISIONI AMMINISTRATIVE E “POVERTÀ” DI
RISORSE DELLA PUBBLICA AMMINISTRAZIONE.

Francesco de Leonardis

1.	La prospettiva d’indagine.	137
2.	Le tre tappe della digitalizzazione, interconnessione e intelligenza artificiale.....	140
3.	I campi di applicazione dell’analisi predittiva.....	144
4.	I vantaggi della “decisione robotica”.....	147
5.	La prima criticità: il diritto alla riservatezza.....	149
6.	La seconda criticità: l’opacità della decisione.....	152
7.	La “povertà” di risorse dell’amministrazione italiana di fronte agli algoritmi.....	156

LA JUSTICE PRÉDICTIVE

Louis Larret-Chahine

1.	Introduction - La réalisation d’une utopie.	161
2.	Définition de la justice prédictive.....	163
3.	Les bénéfices de la justice prédictive.	163
	A. Un outil d’aide à la décision pour les profes- sionnels du droit.	165
	B. Un levier de transparence et de performance du service public de la justice	167
4.	Pour une critique constructive de la justice prédicti- ve.	168
	A. La nécessité d’un choix éthique.	168
	B. La question de la régulation.	170
5.	Conclusion – Pour une université #legalgeek.....	172

LAW AS A SET OF DECISIONS.
ON MERITS AND DANGERS OF LEGAL REALISM THROUGH
THE PRISM OF BIG DATA

Wojciech Zagorski

1.	Introduction.....	175
2.	The Triumph of Legal Realism.	178
3.	The Dangers of Legal Realism.....	181

THE ROLE OF HUMAN JUDGE IN JUDICIAL DECISIONS.
PRELIMINARY REMARKS ON LEGAL INTERPRETATION IN
THE AGE OF ARTIFICIAL INTELLIGENCE.

Laura Vagni

1. The task of judging and the option of an artificial judge.....	185
2. Law and interpretation or law as interpretation.....	191
3. The influence of the human judge on judicial decisions.....	195
4. The role of the judge: a constant tension between human being and being human.....	199

AUTORI

ERMANNO CALZOLAIO

Professore ordinario di Diritto Privato Comparato, Dipartimento di Giurisprudenza, Università di Macerata.

CARLA CANULLO

Professore associato di Filosofia Teoretica, Dipartimento di Studi Umanistici, Università di Macerata.

GARANCE CATTALANO

Professeur de Droit Privé, Faculté de Droit et Gestion, Université de Orléans.

FRANCESCO DE LEONARDIS

Professore ordinario di Diritto Amministrativo, Dipartimento di Giurisprudenza, Università di Macerata.

FRÉDÉRIC DOURNAUX

Professeur de Droit Privé, Faculté de Droit et Gestion, Université de Orléans.

EMANUELE FRONTONI

Professore associato di Sistemi di Elaborazione dell'Informazione, Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche.

NICOLAS HAUPAIS

Professeur de Droit Public, Faculté de Droit et Gestion, Université de Orléans.

LOUIS LARRET-CHAHINE

Fondatore e Direttore Generale di «Predictice».

MARINA PAOLANTI

Ricercatrice di Sistemi di Elaborazione dell'Informazione, Di-

partimento di Ingegneria dell'Informazione, Università Politecnica delle Marche.

STEFANO POLLASTRELLI

Professore ordinario di Diritto della Navigazione, Dipartimento di Giurisprudenza, Università di Macerata.

PIERRE SERRAND

Professeur de Droit Public, Faculté de Droit et Gestion, Université de Orléans.

MONICA STRONATI

Professore associato di Storia del Diritto, Dipartimento di Giurisprudenza, Università di Macerata.

LAURA VAGNI

Professore associato di Diritto Privato Comparato, Dipartimento di Giurisprudenza, Università di Macerata.

WOJCIECH ZAGORSKI

Maître de conférences en Droit Public, Faculté de Droit et Gestion, Université de Orléans.

INTELLIGENZA ARTIFICIALE ED AUTONOMIA DELLA DECISIONE: PROBLEMI E SFIDE

Ermanno Calzolaio

SOMMARIO: 1. Spunti introduttivi. – 2. Le innovazioni tecnologiche e il diritto. – 3. La decisione quale criterio di osservazione del fenomeno.

1. Spunti introduttivi.

Le nuove tecnologie si sviluppano incessantemente e attirano sempre più l'attenzione del giurista, impegnato a confrontarsi con gli scenari inediti che si schiudono con la massiccia utilizzazione degli algoritmi e dell'intelligenza artificiale.

Un primo ostacolo è costituito dalla terminologia utilizzata, che spesso si rivela piuttosto opaca. Le medesime espressioni sono utilizzate con significati e in contesti diversi, con la conseguente possibilità di gravi equivoci nella loro applicazione al linguaggio giuridico¹.

Per cominciare, allora, che cosa si intende per intelligenza artificiale?

Senza alcuna pretesa di completezza, è utile ricordare che l'espressione viene coniata da alcuni studiosi statunitensi già negli anni '50 dello scorso secolo, per designare la tendenza ad "umanizzare" le macchine, dal punto di vista del linguaggio, dei concetti e delle problematiche connesse. Da quel momento in poi, le definizioni si sono moltiplicate e la letteratura è divenuta sterminata, fino a rimettere in discussione la stessa termino-

¹ In questo senso, v. I. GIUFFRIDA - F. LEDERER - N. VERMEYS, *A legal perspective on the trials and tribulations of AI: how artificial intelligence, the Internet of Things, Smart Contracts, and other technologies will affect the Law*, in *Case Western Reserve Law Review*, 2018, pp. 747 ss., in specie a p. 751.

logia utilizzata, sottolineando che in realtà ciò che le macchine sono in grado di fare è lontano dalla funzione cognitiva che caratterizza l'intelligenza umana. Infatti, è ben vero che i sistemi di intelligenza artificiale “apprendono”, ma in modo profondamente diverso dagli esseri umani. Per raggiungere un risultato utilizzano i *proxy*, includendo correlazioni statistiche che derivano dai dati acquisiti, ma un modello di apprendimento automatico non è in grado di comprendere concetti astratti in relazione agli obiettivi che vuole raggiungere².

Sempre a livello introduttivo, va pure ricordato che l'espressione “intelligenza artificiale” designa, in realtà, almeno quattro tecnologie diverse: l'apprendimento (*machine learning*), il riconoscimento del discorso (*speech recognition*), il trattamento automatico del linguaggio (*natural language processing*) e il riconoscimento dell'immagine (*image recognition*).

Tra esse, assume particolare interesse il *machine learning*, che può essere definito come l'abilità di un computer a far evolvere i suoi programmi prendendo in considerazione nuovi dati. Il *machine learning* utilizza i computer per implementare modelli predittivi, che apprendono a partire dai dati già esistenti per prevedere i comportamenti, i risultati e le tendenze future. Pertanto, il *machine learning* riposa e dipende dai dati. Apprende meglio a seconda della quantità dei dati a sua disposizione. Nello stesso tempo, il modo con cui i dati sono utilizzati può porre dei problemi relativamente alla qualità, alla precisione e all'utilità delle informazioni generate.

Un esempio illustra bene i rischi: l'algoritmo utilizzato da Google per rispondere alle domande dell'utente, consultato circa la religione di Barack Obama, ha risposto che l'allora Presidente degli Stati Uniti era musulmano, mentre in realtà è cristiano. L'algoritmo non sbagliava. Aveva semplicemente utilizzato dei dati errati, reperiti nell'universo di internet, senza avere la possibilità di riconoscere che si trattava di una *fake news*, non essendo in grado di distinguere tra notizie buone e cattive.

² R. BONNAFFÉ, *New Technologies and the Law: The Impact of Artificial Intelligence on the Practice of Law*, disponibile al seguente indirizzo: <https://medium.com/@remybonnaffe/new-technologies-and-the-law-the-impact-of-artificial-intelligence-on-the-practice-of-law-c456904688d1>.

2. Le innovazioni tecnologiche e il diritto.

Il giurista si confronta da sempre con le sfide poste dall'innovazione tecnologica. Storicamente, il suo approccio si è fondato sul ricorso a categorie e strumenti interpretativi classici (basta pensare all'analogia) per disciplinare i nuovi fenomeni sul piano legislativo, giurisdizionale e scientifico, ricercando una via mediana tra tradizione e innovazione. Nel XXI secolo potrà operare allo stesso modo o dovrà cambiare la sua *forma mentis* e il suo *modus operandi*? La domanda si impone, perché oggi l'innovazione investe la struttura stessa della conoscenza giuridica. Ci si trova di fronte a quella che si è prestata ad essere letta come una vera e propria rivoluzione cognitiva, destinata a rimettere in discussione categorie e strumenti concettuali consolidati³.

Nel tentativo di cogliere in quali modi il cambiamento tecnologico sta investendo il mondo del diritto, è stata proposta una classificazione articolata su tre livelli⁴. Essa, sebbene formulata con specifico riferimento al funzionamento della giustizia, è però utile a cogliere alcuni aspetti del fenomeno nel suo complesso.

Vi è, anzitutto, una “tecnologia di supporto” (*supportive technology*), per abbracciare tutte le ipotesi in cui le nuove tecnologie rendono possibile l'accesso gratuito a informazioni (anche mediante delle “app”), che facilitano la ricerca di legislazione e giurisprudenza, fino a permettere di consultare un avvocato o trovare assistenza in giudizio⁵. Ciò che caratterizza questo primo livello è, appunto, un supporto, una facilitazione, un miglioramento, non suscettibili di alterare profondamente il mondo del diritto.

Il discorso è diverso quando la tecnologia permette di sostituire alcune attività tipicamente svolte dalla persona, mediante l'automatizzazione di alcuni processi decisionali nel campo amministrativo, contrattuale, di amministrazione della giustizia. Si parla, in questo caso, di tecnologia sostitutiva (*replacement tech-*

³ Cfr. B. DONDERO, *Justice prédictive: la fin de l'aléa judiciaire*, in *Rec. Dalloz*, 2017, p. 532 e A. GARAPON - J. LASSÈGUE, *Justice digitale : révolution graphique et rupture anthropologique*, Paris, 2018, p. 104.

⁴ Cfr. T. SOURDIN, *Justice and technological innovation*, in *Journal of Judicial Administration*, 2015, p. 96.

⁵ http://www.abajournal.com/magazine/article/20_apps_providing_easier_access_to_legal_help.

nology). Con specifico riferimento al mondo della giustizia, ad esempio, sono già in uso strumenti che permettono di organizzare i documenti delle controversie per semplificare e ridurre il tempo di analisi (è quanto avvenuto ad esempio per i contenziosi insorti dopo il crollo del ponte Morandi di Genova, attraverso l'acquisto di un costoso software), oppure che consentono di implementare mezzi di risoluzione alternativa delle controversie (ADR) ⁶, o, ancora, che rendono possibile tenere udienze in teleconferenza ⁷. In tutti questi casi la tecnologia non è più un mero supporto, ma giunge a sostituire almeno un segmento del processo decisionale, inteso in senso ampio.

Questa sostituzione diviene ancor più evidente nel terzo livello, designato come tecnologia dirompente (*disruptive technology*). L'intelligenza artificiale può giungere a creare modelli di decisione che possono essere utilizzati dalla pubblica amministrazione o dal giudice per dirimere una controversia. Solo per fare un esempio, in Messico viene utilizzato un sistema in base al quale il giudice può verificare, in modo automatico, se una domanda di pensione è ammissibile o meno ⁸. Più in generale, è il tema posto dalla giustizia predittiva, di cui oggi sempre più si discute.

⁶ Del resto, il legislatore europeo è intervenuto (Regolamento (UE) n. 524/2013 del Parlamento Europeo e del Consiglio del 21 maggio 2013) per implementare un sistema di risoluzione delle controversie online dei consumatori, mediante l'uso di una piattaforma che facilita le modalità di reclamo e di composizione amichevole, in particolare negli scambi internazionali, in cui non sempre è facile reperire informazioni sul soggetto competente. Questa piattaforma oggi è operativa e impone nuovi obblighi ai professionisti, come quelli di indicare un indirizzo e-mail di contatto sul proprio sito internet (art. 14.5 del Regolamento ODR n. 524/2013), di informare i consumatori circa l'esistenza della piattaforma e la possibilità di ricorrervi per risolvere le controversie, di rendere disponibili le condizioni di vendita sul sito internet, ecc. Questa piattaforma è, in sostanza, un sito internet interattivo, gratuito e multilingue, che svolge le funzioni di uno sportello unico per contattare la controparte per proporre il reclamo e trasmettere il fascicolo, in caso di esito non favorevole, al Mediatore competente. Cfr. T. SOURDIN, *Judge v. Robot, Artificial Intelligence and Judicial Decision-Making*, in *UNSW Law Journal*, 2018, p. 1114, alle pp. 1118 ss.

⁷ Circa l'esperienza australiana, v. E. ROWDEN, *Distributed Courts and Legitimacy: What Do We Lose When We Lose The Courthouse*, in *Law, Culture and the Humanities*, 2018, p. 263.

⁸ Ci si riferisce al sistema *Experitus*, utilizzato in Messico, su cui cfr. D. CARNEIRO E ALTRI, *Online Dispute Resolution: An Artificial Intelligence Perspective*, in *Artificial Intelligence Review*, 2014, p. 211.

3. La decisione quale criterio di osservazione del fenomeno.

Le osservazioni che precedono sono sufficienti per far emergere che uno degli aspetti più rilevanti della diffusione degli impieghi dell'intelligenza artificiale è che con essa sembra incrinarsi, se non proprio spezzarsi, uno dei cardini del pensiero giuridico moderno, costituito dalla possibilità di imputare una condotta ad un soggetto in base alla autonomia del processo decisionale che ne è alla base⁹.

La scelta di concentrarsi sul tema della decisione nasce proprio dal tentativo di individuare una chiave di lettura che consenta di attraversare trasversalmente la complessità e la vastità dei profili connessi allo sviluppo delle nuove tecnologie nel mondo del diritto, cogliendone uno dei nodi più delicati.

I contributi raccolti nel presente Volume ruotano, dunque, attorno a questa area tematica, nella consapevolezza che è quanto mai necessario muovere da un approccio interdisciplinare e aprirsi al dialogo con i cultori delle scienze informatiche, che progettano algoritmi e ne implementano l'impiego¹⁰. Per questa ragione, il saggio introduttivo, redatto da Emanuele Frontoni e Marina Paolanti, specialisti di sistemi di elaborazione delle informazioni, è dedicato ad illustrare il funzionamento degli algoritmi, fornendo alcuni esempi delle concrete applicazioni dell'intelligenza artificiale e della robotica, al fine di mettere a fuoco alcuni nodi problematici posti dall'indipendenza dell'algoritmo dalla decisione umana.

Il tema della sostituzione (anche parziale) con una macchina del soggetto che decide suscita interrogativi profondi, investendo la problematica delle trasformazioni del soggetto. Su questi aspetti si concentra, da una prospettiva filosofica, il saggio di Carla Canullo, muovendo proprio dalla domanda "chi decide?".

⁹ Sul tema della decisione cfr. ad esempio G. RESTA, *Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza*, in *Politica del Diritto*, 2019, pp. 199 ss., nonché G. NOTO LA DIEGA, *Against the dehumanisation of Decision-Making*, in *Journal Intell. Prop. Info. Tech. & Elec. Com.*, 2018, p. 1.

¹⁰ Vengono qui raccolti i primi risultati di una ricerca, dal titolo "Decision and New Technologies" (DANT), che costituisce una delle iniziative di implementazione del progetto del Dipartimento di Giurisprudenza di Macerata, quale "Dipartimento di Eccellenza", finanziato dal MIUR, il cui tema generale è: "Diritto e Innovazione: Europa e Cina di fronte alle sfide della globalizzazione".

Una domanda a cui non è possibile dare una risposta univoca. Ma il fatto che si tratti di una questione aperta è una risorsa, perché essere “davanti a” ogni realtà è la *chance* di “chi decide” e, in tal senso, anche essere davanti all’intelligenza artificiale è uno dei modi in cui tale *chance* accade e si concretizza.

Da questi angoli di osservazione prendono le mosse i saggi più specificamente giuridici, aperti dal contributo di Monica Stronati, che colloca il tema dell’innovazione tecnologica in una prospettiva storica, con particolare riferimento al periodo tra il XIX e il XX secolo, in cui il giurista si è trovato ad affrontare le sfide della rivoluzione industriale, ad esempio rispetto al problema della responsabilità per danni.

Sfide altrettanto delicate connotano, poi, un settore particolare del diritto internazionale, relativo all’utilizzazione delle armi. Sulla diffusione di armi automatiche o semi-automatiche riflette nel suo contributo Nicolas Haupais, rilevando che alla base della de-umanizzazione nel processo di determinazione di un bersaglio e del suo trattamento vi è il tentativo di allontanare gli eventuali responsabili degli atti, di cui sarebbero normalmente chiamati a rispondere sul piano del diritto internazionale.

Venendo poi al diritto privato, la progressiva erosione dell’autonomia privata è evidente anzitutto in materia contrattuale. Basti pensare, con riferimento al diritto dei contratti, agli *smart contracts*, che sono, come è noto, dei protocolli informatici che eseguono operazioni pre-programmate dalle parti. Si tratta di accordi che esistono nella forma di un algoritmo e la cui esecuzione è automatica e indipendente dalla cooperazione del debitore. È l’universo *blockchain*, cioè una tecnologia che consente di registrare contratti senza che sia possibile poi operare qualunque tipo di recesso. All’inizio, questa tecnologia è stata utilizzata prevalentemente nella moneta elettronica (il *bitcoin*), ma ora essa si sta diffondendo in modo esponenziale: in alcuni paesi è sperimentata per la gestione dei registri immobiliari, negli Stati Uniti per la registrazione dei brevetti, e via di seguito¹¹. Anche il legislatore italiano se ne è occupato di recente: nell’ambito di una serie di disposizioni introdotte in materia di sostegno e semplificazione per le imprese e per la pubblica amministrazione, è

¹¹ Per una panoramica delle attuali utilizzazioni cfr. R. HERIAN, *Legal Recognition of Blockchain Registries and Smart Contracts*, 2018 (<https://www.researchgate.net/publication/329715394>).

stata inserita una norma secondo cui gli *smart contracts* soddisfano il requisito della forma scritta, purché le parti interessate siano preventivamente identificate attraverso una procedura le cui caratteristiche verranno fissate con apposito decreto attuativo¹². Basta pensare ai principi classici della teoria del contratto (dovere di cooperazione del debitore, effetto traslativo legato al consenso legittimamente manifestato, ecc.) per rimettere, quanto meno, in discussione la loro utilità per affrontare i problemi che si affacciano rispetto a questi nuovi fenomeni. I saggi di Garance Cattalano e di Frédéric Dournaux si interrogano su tutti questi aspetti, ponendo in luce come le nuove tecnologie sono suscettibili di incidere, a vario titolo, sulla decisione in materia contrattuale.

Oltre al diritto dei contratti, anche quello della responsabilità civile si confronta con la diffusione delle nuove tecnologie. Un profilo di particolare rilievo è costituito dalla guida autonoma (*driverless*) dei veicoli, già in uso e di cui è prevista a breve la commercializzazione su ampia scala, tanto da attirare l'attenzione dei legislatori, che si sono affrettati ad adottare specifiche normative, a livello nazionale, europeo ed internazionale¹³. I principi tradizionali in materia di responsabilità civile sono adeguati per affrontare i problemi che si porranno? In caso di incidente, chi sarà ritenuto responsabile? L'utilizzatore? Il produttore? E in quali limiti? Sarà possibile ritenere responsabile il proprietario anche in assenza di sua colpa? Quali implicazioni si possono immaginare sul piano dogmatico, considerando i fondamenti classici della teoria della responsabilità civile e penale? Su questi profili si concentra lo studio di Stefano Pollastrelli.

Passando al diritto amministrativo, va anzitutto chiarito che il termine "decisione" viene utilizzato, nel linguaggio giuridico, con una pluralità di significati e lo studio di Pierre Serrand si sofferma a ricostruirne un quadro di insieme, esaminando l'esperienza francese. Il saggio di Francesco De Leonardis si concentra, poi, su una analisi del fenomeno della diffusione dell'in-

¹² Art. 8 *ter*, D.L. 14/12/2018, n. 135, convertito nella legge 11 febbraio 2019, n. 12.

¹³ Per un quadro comparativo di sintesi circa le normative sinora adottate, cfr. I. FERRARI, *Analisi comparata in tema di responsabilità civile legata alla circolazione dei veicoli a guida autonoma*, in S. SCAGLIARINI (cur.), *Smart Roads and Driveless Cars: tra diritto, tecnologia, etica pubblica*, Torino, 2019, pp. 97 ss.

telligenza artificiale e dei *big data* nel quadro evolutivo della pubblica amministrazione.

L'ultima serie di saggi si sofferma, infine, sulla decisione giudiziale, poiché la stessa attività giurisdizionale si trova ad essere coinvolta dalla diffusione dell'intelligenza artificiale, con meccanismi sempre più sofisticati non solo di ricerca giuridica, ma anche di guida alla decisione. Louis Larret-Chahine prende in considerazione il fenomeno della giustizia predittiva, esponendone le caratteristiche anche attraverso l'illustrazione di un motore di ricerca (*Prédicitive*) che sta conoscendo una diffusione crescente in Francia. Il saggio di Wojciech Zagorski pone in luce, anche in base ad esempi e statistiche, il ruolo normativo della giurisprudenza, evidenziando fino a che punto le decisioni dei giudici riflettono in larga misura l'approccio individuale del giudice al caso concreto (come da tempo evidenziato dal giusrealismo). La riflessione di Laura Vagni, da ultimo, si concentra sul contributo imprescindibile del giudice come persona alla decisione, sottolineando l'importanza dell'educazione del giurista nell'era dell'intelligenza artificiale.

In conclusione, l'auspicio è che gli studi qui raccolti, nel contesto della diffusione di nuovi strumenti destinati a regolare segmenti sempre più importanti della vita sociale, possano contribuire a stimolare la riflessione, cui certamente è chiamato il giurista del XXI secolo.

AI-BASED DECISION SUPPORT SYSTEM: FROM THEORETICAL BACKGROUND TO REAL WORLD APPLICATIONS

Emanuele Frontoni – Marina Paolanti

TABLE OF CONTENTS: 1. Introduction. – 2. Description of real-world applications. – 2.1 An Infrastructure for decision-making to support neonatal clinical care and research. – 2.2 Decision Support System for assisted reproductive technology. – 2.3 Visual and textual sentiment analysis of social media pictures using deep convolutional neural networks. – 2.4 Deep understanding of shopper behaviours and interactions using computer vision. – 3. Conclusion.

1. Introduction.

In recent years, the use of Information and Communication Technologies (ICTs) has gained increasing importance in our daily lives. Currently, ICTs are used in many fields, including public health, road safety, retail and cultural heritage (CH). Decision support systems (DSSs) are hailed as a possible solution to the onerous cognitive burden currently placed on experts in nearly every domain. All actions and human affairs are the results of a decision-making process. The process of deciding is related to problem-solving and is known as a problem-solving action. Decision-making processes can be varied and may depend on the subjects involved or on the field in which they occur.

A DSS [2] is a software system to support decision-making processes. In other words, it provides support to people who have to make strategic decisions within particularly complex contexts - in which it might be difficult to determine which is the right choice to undertake, or which decision-making strategy should be used to reach a given goal. The terms that

identify DSS software applications and effectively summarise their features are:

Decision - indicates that attention is placed on the decision-making processes and on problems faced in the context of management activities, rather than on transactional or reporting processes;

Support - defines the role of information technologies and the quantitative methods used, deriving from operational research. A DSS, on the whole, constitutes an auxiliary tool and is not a substitute for corporate decision-makers, who must also act on the basis of subjective elements that cannot be codified, such as experience and creativity;

System - underlines the existence of a coherent and integrated context of resources, information technologies tools, users and analysis methodologies.

In the field of DSS study, the decision-making process is represented by the model proposed by H. Simon in the 1960s, which is still considered one of the cornerstones of the decision-making process. This model divides the process into three main phases:

Intelligence: this is the phase in which information is gathered, both from the external and internal environment, to identify and circumscribe a problem to be addressed.

Design: this phase consists of understanding the problem, generating possible solutions and analysing them.

Choice: in this phase the evaluation is made, and therefore a choice is made amongst the alternatives formulated in the previous phase. To arrive at this choice, parameters and indicators are defined so as to allow for comparisons between action plans and forecasts of the consequences of the various options.

More recently, many DSS experts are basing their systems on Artificial Intelligence (AI), aiming to replace experienced decision-makers by emulating their behaviours [18]. AI consists of introducing into a computer's memory both the elements related to a certain phenomenon, and the rules relating to what is normally meant by knowledge of a given phenomenon. As noted, AI plays an important role in DSS—especially in relation to the modelling of behaviours and decisions that allow for an increasingly realistic analysis,

which aims to offer a predictive or inferential result in a given area in which it is relevant to decision-making.

AI is a partial reproduction of human intellectual activity (with particular regard to the processes of learning, of recognition and of choice). It is realised through the elaboration of ideal models, or, concretely, by developing machines for this purpose—mainly electronic computers. AI can be also defined as the science that proposes to develop intelligent machines [16].

These theories are connected to a vision of the human mind as a programme: applying the data of stimuli (the input), the human mind produces a certain body behaviour (the output). The brain can be seen as the parallel hardware (made up of neurons and connections among them) on which this programme is performed. So, if we had sufficient computational resources, a computer could simulate a reasoning human and a computer might have the same capabilities as a human mind. Said computer would not necessarily simulate the human brain, but only human reasoning processes. Its hardware could thus differ from that which would be directly obtained by simulating the brain.

The ELIZA program (1965) [20] was a milestone in AI application. This programme simulated the interaction between a Rogerian therapist and a patient. Many patients thought they were talking with a real therapist, and not with a machine. Yet the phrases ELIZA used with patients were not the result of complex reasoning; they were merely the result of matching various keywords in the phrases spoken by the patients. So, although ELIZA passed the Turing test [19], it was difficult to declare this programme as intelligent. Another AI milestone has been reached with the development of machines that are capable of playing certain games at the level of the world's best human players. The most striking example is Deep Blue (1997) [7], which succeeded in beating the world chess champion, Garry Kasparov.

Today's main areas of AI research can be grouped into the following topics:

- Multi-agent systems: how problems can be described and solved relating to collaboration or competition among multiple agents;
- Constraints and research: how one can model problems whose solutions must comply with restrictions

(constraints); and how one can efficiently find their optimal solution;

- Representation of knowledge: how we can model the knowledge in a certain field of application, and how it can be used to act intelligently in this field;
- Machine learning: how machines can learn new and better behaviours from experience;
- Natural language processing: how machines can understand and reason about a text using natural language;
- Scheduling and planning: how to schedule or plan various activities in an optimal way in the presence of limited resources;
- Robotics and vision: how to build an adaptive and flexible robot that knows how to see and perceive the external world, and knows how to decide the best behaviour to adopt based on this perception of the world;
- Management of uncertainty: how good decisions can be made even if some data are missing or vague;
- Web: how to apply AI techniques to the web world.

Machine learning [13] is a technique that allows one or more mathematical models to be constructed, once the correlations among large masses of data have been identified.

Unlike ‘expert systems’, which sought to model management rules based on human reasoning and apply them to incoming data, the machine independently searches for correlations (with or without the supervision of a human being). The machine then deduces rules, creating one or more mathematical models, from the huge amount of data provided. These models allow for the creation of learning algorithms to solve a specific problem.

Deep Learning [9] is an area of Machine Learning that is based on a particular type of data learning. It is characterised by the effort to create a multi-level machine learning model, in which the deeper levels take the outputs of the previous levels as inputs, transforming them and abstracting them more and more. This insight into the levels of learning gives the field its name and is inspired by the way the brain processes information and learns, responding to external stimuli. Just as the brain learns by trying and activates new neurons by learning from experience, so do the architectures responsible for Deep Learning, in which the extraction stages are modified

based on the information received at the input. The development of Deep Learning has taken place consequently and in parallel with the study of AI, and in particular with the study of neural networks. After its beginnings in the 1950s and especially in the 1980s, this area has developed substantially, thanks to Geoff Hinton and the Machine Learning specialists who collaborated with him. In those years, the computer technology was not sufficiently advanced to allow a real improvement in this direction. So we've had to wait until more recent years to see, thanks to the availability of data and computing power, still more significant developments.

Deep Learning combines increasingly powerful computers with special neural networks to understand the patterns found in large data volumes. Deep learning techniques are currently state of the art, due to their ability to identify objects in images and words in sounds. Researchers are now trying to apply these successes in recognising models to more complex tasks—such as automatic language translation and medical diagnosis—as well as to many other important areas, both in the social and business realms.

2. Description of real-world applications.

In this section, we describe various applications based on AI. Four different case studies are specifically explored: medicine, biology, social media intelligence and retail. At the core of these applications are AI algorithms used to develop DSS for each of the challenging domains discussed below.

2.1 An Infrastructure for decision-making to support neonatal clinical care and research.

Technological advances in neonatal intensive care have been the major focus of neonatal clinical research, as has their impact on babies' short- and long-term survival and health outcomes. Neonatal mortality reached an unsurpassable minimum by the late 1990s. However, this has been associated with high rates of neonatal morbidity and neurodevelopmental impairment. Every year there are more than 15 million preterm births worldwide—that is, more than 1 in 10

babies—and the number of cases continues to rise. Prematurity is the leading cause of death in newborns (babies in their first four weeks of life). It is also now the second leading cause of death, after pneumonia, in children under five years of age.

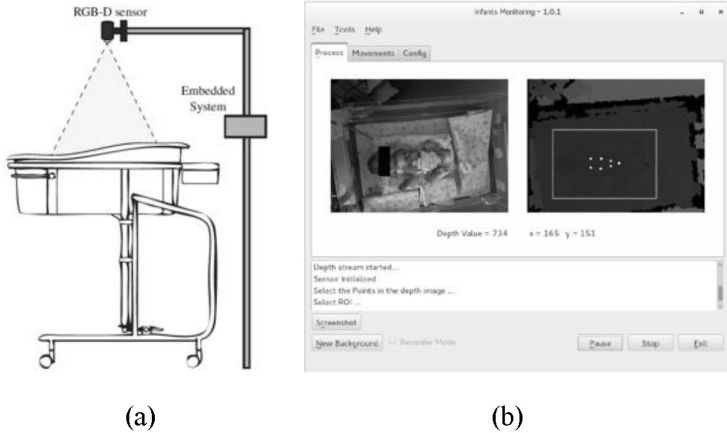


Figure 1: (a) System configuration scheme and (b) application interface

Neonatal intensive care units in hospitals allow medical professionals with specialised training to care for preterm newborns with serious health problems. The vital functions of a baby in an incubator must be continually monitored. Consequently, each baby is connected to multiple probes and devices that generate alarms whenever the limit levels set by the neonatal intensive care unit are exceeded.

The hospital ward taken into consideration for our research is the Neonatal Intensive Care Unit (NICU) of the Women's and Children's Hospital 'G. Salesi' in Ancona. In this NICU, incubators are surrounded by numerous devices for monitoring, diagnosing and treating diseases. The devices provide an enormous amount of data, which is displayed only on monitors. A medical staff member periodically transcribes these data, creating a paper medical record. Yet—to ensure that important details for patient care have not been overlooked—doctors need the data to be collected automatically from various devices. They themselves are aware that the automation of this process can facilitate and improve the implementation of their daily clinical procedures. It can also contribute to the development of

innovative systems, methods and algorithms for the study of some important health problems related to premature babies.

To offer optimal care, it's necessary to achieve the highest possible level of correlation and integration among all the physiological data collected by the instrumentation that is connected to the baby. A critical clinical situation is often caused by a combination of pathologies. It is thus crucial to study the interactions among different physiological parameters.

An architecture has been developed and tested, integrating three different types of tools:

- Patient monitor;
- Bilirubinometer;
- Transcutaneous bilirubinometer.

Medical communication protocols were used to enable communication and data transfer between the child monitoring devices mentioned above and the unique database of the platform. Next, a data collection process was performed from a different instrumentation—standardising the data and forwarding it to the Cloud through the import of Web services. A system built with these features guarantees the interoperability property of the collected data, allowing for the online sharing of data.

For the visualisation of the data, a web interface was designed and developed on the basis of the specifications provided by the healthcare personnel and by the medical guidelines for visualising the data.

An algorithm was also developed to support the diagnosis of neonatal jaundice and the management of therapies to treat the disease. It consists of an automated, evidence-based DSS for assistance, based on guidelines to facilitate the implementation of clinical practice guidelines for neonatal jaundice.

Another proposed algorithm supports the physician in the administration of nutritional fluids. Still another allows for the automation of a clinical test used as follow-up to evaluate the development of preterm infants. An algorithm used to analyse the respiratory rate of preterm newborns has been implemented; this is able to activate an alarm signal when the respiratory rate values leave the physiological range. Through this method—which uses computerised vision techniques and machine learning algorithms—the baby's movements can be detected in real time from the depth flow and can be extracted

from the sequence of depth images, collected by a non-invasive and non-contact RGB-D sensor.

2.2 Decision Support System for assisted reproductive technology.

Reproductive technologies are becoming increasingly important in the medical field. Every year, 10-15% of couples experience problems related to infertility and a growing number of them seek the help of ART (Assisted Reproductive Technology) laboratories. Assisted reproduction technologies have the main objective of obtaining a large number of competent oocytes, which have the ability to be fertilised and to develop as embryos. Currently, in the ART implant, there is no criterion for choosing the oocyte. In the past two decades, even amidst other improvements, the success rates for procedures that treat infertility and lead to live births remain far less than optimal.

Identifying an oocyte of better quality is vital, as it will increase overall pregnancy rates. The accurate prediction of embryos' viability, and their successful implantation, will also reduce the risk of miscarriage and its complications. In this way, the choice of a non-viable oocyte will be reduced avoiding even in repeated multiple gestations.

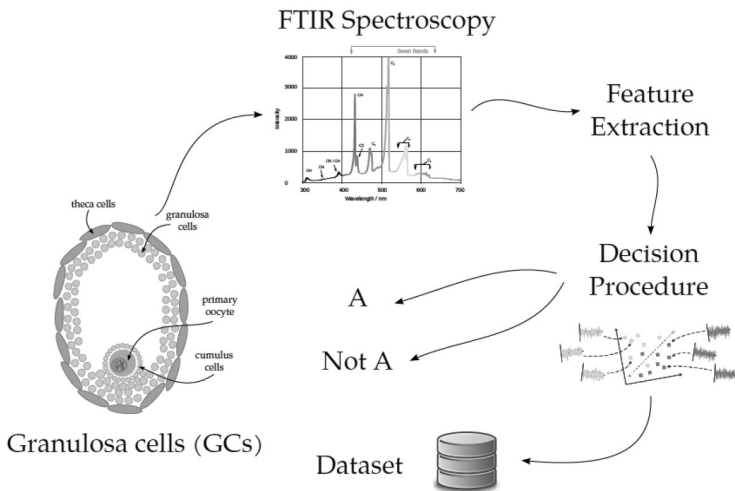


Figure 2: General Schema

Fourier transform infrared microspectroscopy (FTIR) is applied to Granulosa (GC) cells collected together with oocytes during oocyte aspiration, as usually occurs in ART, and spectral biomarkers were selected by multivariate statistical analysis. The GC samples were classified into 2 groups:

- Group A, GC from good quality oocytes;
- Group NON A, GC from poor quality oocytes.

The spectral biomarkers of the two groups described above were used to create a robust biological reference data set (BRD), which was developed retrospectively by analysing GC from known quality oocytes validated by clinical outcomes obtained during the ART procedure of the corresponding oocyte.

Our proposal was to use machine learning algorithms to successfully predict the quality of oocytes for a pregnancy, considering the lack of criteria in ART and considering our previous experience in developing the system of data standardisation in medicine and sharing safe between them. Our research also concerned the comparison of the performances obtained considering 4 classifiers, such as Support vector Machine (SVM) [4] and k-Nearest Neighbour (kNN) [5], Decision Tree (DT) [14] and Random Forest (RF) [3].

The main contribution of our research was: i) the first machine learning approach (ML) developed to determine and classify the quality of oocytes; ii) the introduction of a significant set of bio-characteristics, able to achieve the relative classification performances; iii) the first comparison between different classes, to evaluate how different ML algorithms are able to provide significant results in this field.

2.3 Visual and textual sentiment analysis of social media pictures using deep convolutional neural networks.

The advent of social media now allows anyone to easily create and share ideas, opinions and content with millions of other people around the world. In recent years, we have witnessed the explosive popularity of image-sharing services that not only reflect people's social life, but also express their opinions on disparate topics. These images represent a rich source of knowledge for companies to understand consumers' opinions, yet their multitude makes a manual approach

complex and increases the attractiveness of automatic sentiment analysis.

Automatic algorithms for understanding the feeling or sentiment linked to an image are fundamental to understanding users' behaviour. They are widely applicable to many sectors, from trade to tourism and beyond.

Several researchers have now begun to propose solutions for sentiment analysis of visual content. However, many social images include not just visual but textual elements. To estimate the overall sentiment of an image, it is essential not only to judge the feeling of the visual elements, but also to understand the meaning of the text.

This work proposes a method to estimate the general feeling of an image, based on both visual and textual information. A machine learning classifier based on visual and textual functions, extracted from two specially trained deep convolutional neural networks (DCNNs), was designed to classify the sentiment of an image as positive, negative or neutral.

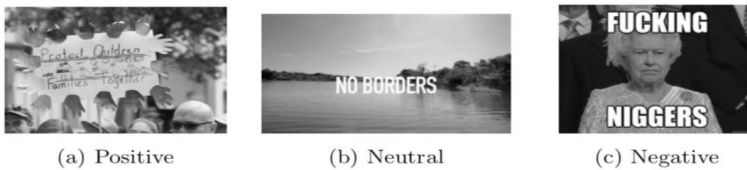


Figure 3: Social Media Pictures from the ‘Social Media PictureS News-related (SIMPSON) Dataset’. (a) is an example of a picture with overall positive sentiment; (b) represents an image with overall neutral sentiment; (c) is a picture with overall negative sentiment.

The visual functionality extractor is based on the VGG16 network architecture [17] and is trained by setting up a model preset on the ImageNet dataset [8]. While the visual element extractor is applied to the entire image, the text element extractor detects and recognises the text before extracting the elements. The text functionality extractor is based on the DCNN architecture proposed by [21] and a fine-tuning model is created that has previously been trained on synthesised social media images. Based on these features, the performances of six classifiers were compared: kNearest Neighbours (kNN), Support Vector Machine (SVM), Decision

Tree (DT), Random Forest (RF), Naive Bayes (NB) [15] and Artificial Neural Network (ANN) [12].

To test our system, we created a new dataset of images extracted from Instagram and called “SoCial Media Pictures News-related” (SIMPSoN) Dataset”. This dataset includes 9247 images containing visual and textual elements labelled as:

- 3446 images with positive sentiment;
- 2903 images with neutral sentiment;
- 2898 images with negative sentiment.

Unlike many existing datasets, the real sentiment was manually estimated by human annotators, and not automatically judged by the associated texts or hashtags, thus providing a more precise dataset. The SIMPSoN Dataset is publicly available¹ for research purposes.

2.4 Deep understanding of shopper behaviours and interactions using computer vision.

Understanding how consumers behave in shopping environments is very important for retailers. Many recent studies have focused on monitoring how shoppers move around in the retail space and interact with products. There are many problems related to these aspects: occlusions, changes in appearance, dynamic and complex backgrounds. The sensors most used for this activity are RGB-D cameras, for their convenience, reliability and availability. Accuracy and effectiveness in the use of depth cameras has been analysed in several studies. While the retail environment has several favourable aspects for computer vision, the large number and diversity of products and the difficulty in correctly interpreting buyers’ movements still make it challenging to accurately measure buyer behaviour.

¹ <http://vrai.dii.univpm.it/content/simpson-dataset>

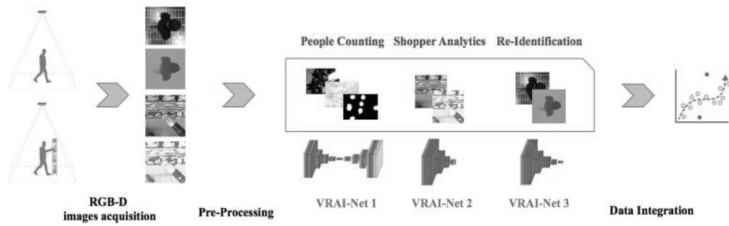


Figure 4: VRAI framework for the deep understanding of shoppers' behaviour.

Recently, feature-based machine learning tools have been developed to learn the skills of shoppers in smart retail environments. Each application uses RGB-D cameras in a top view configuration and provides large volumes of multidimensional data [10,11]. These data are analysed with the aim of examining the attraction, attention and action of consumers.

In this research, we presented a new VRAI² deep learning framework (VRAI) with the aim of improving existing retail applications such as people counting, shopper analytics and re-identification. This process is based on a two-year experience that measured 10.4 million shoppers and about 3 million interactions in stores in Italy, China, Indonesia and the US. In order to conduct a comprehensive evaluation, it is critical to collect representative datasets. While much progress has been made in recent years regarding efforts in sharing codes and datasets, it is of great importance to develop libraries and benchmarks to gauge state-of-the-art datasets.

The framework developed has required three dataset learning methods: Top-View Heads data set (TVHead) to count people³, Hands dataset (HaDa)⁴ and Top-View Person Re- data set Identification2⁵ (TVPR2).

The proposed research starts from the idea of collecting appropriate datasets from real scenarios, with the aim of

² This name of the framework is connected to the Vision Robotics and Artificial Intelligence (VRAI) research group of Università Politecnica delle Marche, to which all authors belong.

³ <http://vrai.dii.univpm.it/tvheads-dataset>

⁴ <http://vrai.dii.univpm.it/content/hada-dataset>

⁵ <http://vrai.dii.univpm.it/tvheads-dataset>

providing a deep learning method with three simultaneous Convolutional Neural Networks (CNNs) that process the same frame to: i) segment and count the people with whom to count high precision even in crowded environments); ii) measure and classify the interactions between buyers and shelves as positive, negative or neutral actions; iii) perform a new identification of contemporary buyers.

3. Conclusion.

The implementation and development of Decision Support Systems (DSS) to support intelligent decision-making has grown in importance as a research field, in recent years. Due to the increased complexity of decision-making, the process now often requires active, intelligent involvement on the part of both user and computer. Artificial intelligence techniques are used in the development of systems that help humans with this decision-making process.

Machine learning is one technique to construct algorithms that can learn from past events and predict unknown events. The result of a learning algorithm is represented by a predictive model, which, when applied to a set of data, allows predictions to be made. These models can be represented with different formats, introducing another level of difficulty into the system. The challenge is to create a decision support system that uses a knowledge base developed via machine learning techniques.

In this study, we present our emerging research into the application of AI to decision support focused within four different environments. We introduced smart, adaptive systems that can modify themselves to solve complex problems in different areas of research.

References

- Alpaydin, E. (2009). *Introduction to machine learning*. MIT press.
- Arnott, D., & Pervan, G. (2005). A critical analysis of decision support systems research. *Journal of information technology*, 20(2), 67-87.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.

Cover, T., & Hart, P. (1967). Nearest neighbour pattern classification. *IEEE transactions on information theory*, 13(1), 21-27.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

Hsu, F. H. (2004). *Behind Deep Blue: Building the computer that defeated the world chess champion*. Princeton University Press.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436.

Liciotti, D., Contigiani, M., Frontoni, E., Mancini, A., Zingaretti, P., & Placidi, V. (2014, August). Shopper analytics: A customer activity recognition system using a distributed rgb-d camera network. In *International workshop on video analytics for audience measurement in retail and digital signage* (pp. 146-157). Springer, Cham.

Liciotti, D., Paolanti, M., Pietrini, R., Frontoni, E., & Zingaretti, P. (2018, August). Convolutional networks for semantic heads segmentation using top-view depth data in crowded environment. In *2018 24th International Conference on Pattern Recognition (ICPR)* (pp. 1384-1389). IEEE.

Lippmann, R. P. (1987). An introduction to computing with neural nets. *IEEE ASSP magazine*, 4(2), 4-22.

Mitchell, T. M. (1997). *Machine Learning*. New York: McGraw-Hill. ISBN: 978-0-07-042807-2.

Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1(1), 81-106.

Rish, I. (2001, August). An empirical study of the naive Bayes classifier. In *IJCAI 2001 workshop on empirical methods in artificial intelligence* (Vol. 3, No. 22, pp. 41-46).

Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited.

Simonyan, K., & Zisserman, A. (2014). Very deep

convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Turban, E., & Watkins, P. R. (1986). Integrating expert systems and decision support systems. *Mis Quarterly*, 121-136.

Turing, A. M. (2009). Computing machinery and intelligence. In *Parsing the Turing Test* (pp. 23-65). Springer, Dordrecht.

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45.

Zhang, X., Zhao, J., & LeCun, Y. (2015). Character-level convolutional networks for text classification. In *Advances in neural information processing systems* (pp. 649-657).

CHI DECIDE? INTELLIGENZA ARTIFICIALE E TRASFORMAZIONI DEL SOGGETTO NELLA RIFLESSIONE FILOSOFICA

Carla Canullo

SOMMARIO: 1. La filosofia moderna, la nascita del soggetto e il sogno della macchina. – 2. La *contrée des philosophes*. – 3. L'apparire inatteso della volontà. - 4. La sfida della libertà – 5. La libertà della decisione.

1. La filosofia moderna, la nascita del soggetto e il sogno della macchina.

La filosofia moderna nasce insieme all'interesse per l'“artificiale”. Descartes, nella Quinta parte del *Discours de la méthode*, descrivendo il corpo come una macchina perfetta, scrive che tale paragone «non sarà strano per coloro che sanno quanti automi (*automates*) o macchine che muovono se stesse l'industria degli uomini è capace di produrre, impiegando pochissimi elementi»¹. Pochi anni dopo, Pascal e Leibnitz inventeranno le prime macchine per rendere più efficiente il calcolo nella consapevolezza, che era già cartesiana, che questi strumenti non condideranno mai con l'uomo l'uso della parola. Certo, oggi la situazione è cambiata ed è imparagonabile con quanto avveniva nei secoli XVII e XVIII. Tuttavia, non è da escludere che, dato il loro interesse per l'innovazione tecnica e tecnologica, gli stessi Descartes, Pascal e Leibnitz sarebbero interessati sia a quanto oggi sta accadendo nell'ambito dell'intelligenza artificiale, sia alle implicazioni che ciò reca con sé.

Tornando a Descartes, è nota l'idea del soggetto moderno che egli inaugura, ossia l'idea di un “io” che è irriducibile a ogni

¹ R. DESCARTES, *Discours de la méthode*, Paris, 1992, p. 120, trad. nostra.

dubbio e perciò si scopre come fondamento di ogni sapere. Se, infatti, di tutto posso dubitare tranne che di me che sto dubitando, affermare “je suis, j'existe” e “je pense, donc je suis” (o nella versione latina “cogito ergo sum”) equivarrà a fare del soggetto il fondamento della consapevolezza di sé e, a partire da ciò, di ogni altro sapere. Il che, tuttavia, rappresenta soltanto il primo aspetto della questione. Infatti, una volta che il cammino è stato aperto, anche la concezione della verità si destina al cambiamento, ché l'idea di una verità come *certezza* di cui (l') *esprit/mens* sarebbe il garante prende il posto del paradigma medievale secondo cui «veritas est adaequatio intellectus ad rem» e/o «adaequatio intellectus et rei».

Ora, proprio questo soggetto (al contempo fondamento di ogni conoscenza possibile e garante della verità), è stato messo in crisi nella filosofia del XX secolo e XXI secolo che ne ha criticato il suo essere legislatore universale astratto e disincarnato. Un soggetto che, come ebbe a dire Paul Ricœur coniando una celebre formula, “n'est, à vrai dire, personne”. E tale “io” o “soggetto” è “nessuno” perché, in luogo di essere individuo in carne e ossa, è funzione trascendentale di ogni conoscenza.

Per uscire da questa funzionalità neutra, Martin Heidegger (e con lui molti altri filosofi) hanno sostituito al soggetto e all' “io” il pronome “chi”, e ciò allo scopo di pensare un soggetto di cui non si parli in astratto ma *in concreto*. “Chi”, infatti, non è un pronome personale diretto ma, in quanto pronome relativo, interrogativo o indefinito, indica sempre “qualcuno che” compie un'azione. Questo *glissement* dall'astratto al concreto è stato declinato da Ricœur così: «(Riflettere sul “chi”) vuol dire rispondere alla domanda: *chi* ha fatto questa azione? *Chi* ne è l'agente, l'autore? [...] Rispondere alla domanda “chi?” [...] vuol dire raccontare la storia di una vita. La storia raccontata dice il *chi* dell'azione»². Ora, niente vieta di prolungare la lista delle azioni che Ricœur elenca aggiungendovi la domanda «*chi* decide?», ché anche questa domanda ci condurrebbe 1- ad abbandonare l'universalità (vuota) del soggetto 2- al fine di coglierne la concretezza 3- facendo accadere al soggetto stesso una trasformazione che, forse, è ben più radicale di quella che l'intelligenza artificiale ha provocato.

² P. RICŒUR, *Il tempo raccontato*, trad. G. Grampa, Milano, 1998, p. 375.

2. La *contrée des philosophes*.

L'interesse della filosofia per l'intelligenza artificiale ha dato vita a un dibattito che, grosso modo, si svolge attorno a tre punti principali: 1- che cosa sono la mente e gli stati mentali, 2- di che genere è il rapporto tra mente e cervello e 3- se esista o meno uno stato autonomo e originale di coscienza. Questi tre punti sono stati messi a fuoco dopo l'"evento" che ha avviato tali ricerche, ossia l'invenzione della "macchina" di Alan Turing, nella quale è stato fin da subito individuato uno strumento atto a esplicitare il funzionamento della mente umana. Tornare a questo evento inaugurale è necessario per chiarire perché la filosofia si interroga *meno* sui singoli aspetti e dibattiti in cui l'intelligenza artificiale è implicata e *più* su tutto ciò che è computazione, algoritmo, calcolo computazionale.

Sulla scia di Turing, è stato detto, l'interesse della filosofia si è rivolto a ciò che può chiarire il funzionamento della mente. Hilary Putnam, ad esempio, ha teorizzato l'analogia tra mente e *computer*, analogia che permetterebbe «di cogliere che parole come "pensare", "ricordarsi di", "comprendere", "vedere", (termini che indicano le nostre attività mentali) non dipendono da una sostanza spirituale»³ ma sono azioni che anche le macchine compiono. Tuttavia, prima di Putnam è stato il già citato Alan Turing, con quello che è chiamato "gioco di Turing", ad affermare che "il computer pensa".

Il gioco che ha condotto il matematico e filosofo inglese per dimostrare questa constatazione consiste in tre partecipanti, un uomo A, una donna B (entrambi devono trovarsi nella stessa stanza) e un terzo partecipante C che si trova in una stanza diversa. Utilizzando il tramite di una macchina che trascrive, oggi connettendo in rete due computer, il terzo interlocutore deve individuare il sesso di chi risponde alle domande senza mai ascoltare la voce di nessuno dei due interlocutori. Una variante del gioco prevede la sostituzione del partecipante A o B con una macchina (senza che il terzo interlocutore ne sia al corrente). L'esperimento ha mostrato che il partecipante C, dovendo indovinare se le risposte alle sue domande vengono da un uomo, una donna o una macchina, non ha elementi per sciogliere la questio-

³ H. PUTNAM, *Mind, Language and Reality: Philosophical Papers*, Cambridge Mass, 1975, p. 291, trad. nostra.

ne perché le risposte non presentano differenze (né di genere, di altro tipo). Tenendo dunque conto di questo esperimento, Putnam ha sostenuto che «le immagini mentali e le parole pensate non rappresentano in modo intrinseco ciò di cui sono rappresentazioni»⁴, motivo per cui esse possono essere ricondotte indifferente alle persone o alla macchina. Le neuroscienze, che si sono sviluppate proprio a partire da questi presupposti, offrirebbero altre prove a supporto della tesi.

Secondo questa ipotesi di lettura, non esisterebbe un “chi” centro di decisioni ma l’analogia con l’intelligenza artificiale mostrerebbe la sostanziale assimilabilità di ciò che chiamiamo “libertà” e di ogni attività caratterizzante “chi” esiste, si tratti di stati mentali o di proprietà cerebrali.

Contestando la linea Turing-Putnam, John Searle ha sostenuto che l’intelligenza artificiale non dispone di atti autenticamente soggettivi che, nel migliore dei casi, sono simulati da un calcolatore elettronico. Riprodurre atteggiamenti non significa esserne l’autore e il soggetto responsabile. Gli stessi calcoli che un’intelligenza artificiale compie tramite gli algoritmi non esistono in natura ma sono il risultato di una coscienza che li ha programmati. Secondo Searle, «la coscienza è un vero e proprio fatto biologico, mentre il calcolo è una proprietà da attribuire al sistema elettronico»⁵. Ancora, «per me, scrive Searle, gli stati mentali sono altrettanto reali quanto lo sono gli altri fenomeni biologici, reali come lo è [...] la digestione». In altri termini, le caratteristiche della coscienza, la quale è diversa dal cervello ed è irriducibile agli stati fisici, non possono essere estese all’intelligenza artificiale.

Ora, soltanto in apparenza queste osservazioni sono lontane dal nostro contesto che s’interroga sul “chi” della decisione. Infatti, tra i problemi che Searle pone c’è quello che chiede se una macchina, certamente capace di manipolare simboli, sia capace di comprenderne il significato. La risposta del filosofo è negativa, poiché saper riprodurre simboli non significa che li si comprenda. E così come l’universo simbolico è un tratto esclusivamente appartenente all’uomo, anche la decisione lo è: una cosa è

⁴ H. PUTNAM, *Reason, Truth and History*, Cambridge Mass, 1981, p. 19, trad. nostra.

⁵ J. SEARLE, *Intentionality. An Essay in the Philosophy of Mind*, Cambridge Mass, 1983, p. 312, trad. nostra.

dare conto di decisioni, tutt'altra prenderle ed esserne responsabili.

Volendo sintetizzare questo dibattito (ormai classico) sull'intelligenza artificiale per ricondurlo nel quadro delle questioni che ci interessano, diremo che esso si interroga sulla questione del "chi" interrogando la natura della mente e chiedendo se il suo naturalismo sia riproducibile o no dalla macchina. Ma, appunto, tanto nella linea Putnam/Dennett quanto in quella di Searle, a essere in questione è il naturalismo, motivo per cui, a nostro avviso, nessun passo in avanti è compiuto nella questione "chi decide". D'altronde, per quanto la "macchina" sia assimilabile alla mente umana quanto al funzionamento, essa è in ogni caso priva di intenzioni, né potrebbe comprendere il linguaggio simbolico o porre autonomamente a se stessa degli scopi.

Lo ha colto Henri Atlan quando chiedeva: «È possibile concepire una macchina nella quale gli obiettivi da raggiungere, i compiti da eseguire, non siano imposti dall'esterno, ma *prodotti dalla macchina stessa?* [...] Può una macchina essere intenzionale? [...] "Funzioni" e "finalità" non verranno utilizzate, allo scopo di spiegare il comportamento del sistema. Al contrario, sono proprio *queste* le nozioni che andranno spiegate, utilizzando soltanto quelle spiegazioni di tipo causale che la dinamica delle reti di automi è in grado di fornire, per rendere conto dell'emergenza delle funzioni e dei fini»⁶.

Giovanni Iorio Giannoli ha portato avanti la via aperta da Atlan, sostenendo tuttavia che potremmo assumere che gli studiosi dei sistemi complessi sappiano simulare l'emergenza di certe *strutture* (ad esempio, regioni di una rete) «che si comportino in maniera *coerente* e che siano *stabili*, cioè poco sensibili alle perturbazioni, e di certe *funzioni* (ad esempio, regioni o elementi della rete che *riconoscano* certe classi di perturbazioni, distinguendole nettamente da altre)». Inoltre, se supponiamo di poter assegnare a un sistema del genere un "modulo di auto-osservazione", potremmo giungere a creare una sorta di «*memoria*, il cui compito è quello di immagazzinare la *sequenza degli stati* che il sistema ha attraversato, prima di svolgere con *successo* una certa funzione». In tal modo potremmo, almeno in teoria, ri-

⁶ H. ATLAN, *Intentional Self-organisation. Emergence and Reduction: Towards a Physical Theory of Intentionality*, in *Thesis Eleven*, 52, 1998, trad. nostra.

solvere il problema perché «con questi pochi ingredienti è possibile simulare la *capacità di fare progetti*. In una particolare accezione, “progettare” significa infatti “capacità di prevedere gli effetti di un certo comportamento”; ma tale capacità è garantita dal fatto che si abbia il *ricordo di successi già prodotti* da un comportamento analogo, in una precedente occasione. Noi progettiamo, perché *sappiamo quale sarà l’esito di certe azioni*, sulla base di attività precedenti, fissate nella nostra memoria. Ad esempio, secondo questo schema, un primate potrebbe essere in grado di riprodurre “intenzionalmente” un certo strumento, quando questo strumento sia stato prodotto la prima volta per caso ed il processo di produzione sia stato memorizzato»⁷.

Alla domanda se la macchina sarà un giorno capace di assumere decisioni autonome, Iorio Giannoli risponderebbe che al momento ciò non è possibile, sebbene lo scopo del programmatore sia quello di avvicinarsi il più possibile alla progettazione di azioni basate su stati memorizzati. Un esito, questo, verso cui l’intelligenza artificiale tende e di fronte al quale la *contrée des philosophes* mostra non poca curiosità e attenzione: fino a che punto, infatti, l’uomo potrà conservare la propria originale eccezionalità?

3. L’apparire inatteso della volontà.

Giunti a questo punto, alla domanda se l’intelligenza artificiale abbia trasformato il soggetto, potremmo rispondere che certamente ne ha *trasformato* il dibattito conducendo a chiedere *chi* sia il soggetto che conosce e se la conoscenza, o addirittura l’uomo, abbiano una propria originalità. L’intelligenza artificiale “risponde” a tali questioni *eliminando* il residuo mistico e spirituale dell’intelligenza del soggetto e *rendendo* possibile la comprensione del suo funzionamento. In ciò essa è a pieno titolo argomento di cui si occupa la filosofia, la quale fin dalle sue origini si è interrogata sulla possibilità di rendere comprensibile e migliorabile il rapporto dell’uomo con il mondo che lo circonda. Facendolo, il pensiero filosofico ha formulato modelli che hanno

⁷ G. IORIO GIANNOLI, *Intelligenza artificiale e filosofia*, cit. da https://www.academia.edu/3575747/Intelligenza_artificiale_e_filosofia, p. 16, consultato in data 20/04/2019.

contribuito e contribuiscono a tutt'oggi allo sviluppo delle scienze e *anche* dell'intelligenza artificiale.

Tale contributo, tuttavia, sembra evidenziarsi *meno* nella *contrée des philosophes* e *più* nell'ambito della ricerca sull'intelligenza artificiale stessa. Infatti, l'esplicazione più interessante di tale apporto è a nostro avviso stata data proprio da una delle figure di spicco dello sviluppo della robotica e dell'*Artificial Intelligence*, John McCarthy. Questi, nel 1996 nell'articolo *What has AI in Common with Philosophy?*⁸ sosteneva l'importanza del ricorso dell'intelligenza artificiale alla filosofia, e ciò per diverse ragioni legate soprattutto alla robotica e alla programmazione dei robot, giacché «robots that do what they think their owners want will have to reason about wants».

Il *robot* di cui parla McCarthy incuriosirebbe certamente Descartes, interessato all'*automate*. Tuttavia, lo interesserebbe non perché pone problemi di filosofia della mente (nei quali, come abbiamo visto, la filosofia cartesiana è perfettamente a suo agio avendoli *inventati*), ma lo interrogherebbe in merito a ciò che la sua filosofia *non ha* e che ha bisogno di ricevere da altro che dalla filosofia, ossia il dibattito sulla volontà. Infatti è ancora McCarthy a scrivere che «an attitude toward the free will problem needs to be built into robots in which the robot can regard itself as having choices to make, i.e. as having free will», e la decisione è questione di *free will*. Descartes, dicevamo, sarebbe interessato per un semplice motivo: perché mentre, da un lato, egli apre la via del soggetto, dall'altro lato lascia ogni questione legata alla morale alla provvisorietà (parlando come noto di “morale provvisoria”) e si limita a trattare della volontà là dove parla dell'errore: noi sbagliamo perché la nostra volontà decide anche su cose che non ha conosciuto. Ciò accade perché mentre sull'*esprit/mens* possiamo raggiungere la massima delle certezze, sulla volontà e le decisioni che essa guida non possiamo raggiungere nessuna verità certa che sia condivisibile da tutti. Perciò riflettere sul libero volere facendolo attraverso un'intelligenza di fatto prodotta, significherebbe contribuire al chiarimento di quanto nessuno è mai riuscito a chiarire, ossia al chiarimento del perché liberamente decidiamo e vogliamo qualcosa. Facilmente

⁸ J. MACCARTHY, *What has AI in Common with Philosophy?*, consultabile al sito <http://www-formal.stanford.edu/jmc/aiphil.pdf> (ultimo accesso in data 20/07/2019).

si vedrà che questo tema è fino a ora mancato dal dibattito filosofico che è stato esposto. In tale dibattito, il tema più prossimo alla decisione libera è quello dell'intenzionalità.

Con questo termine è indicato il riferimento interno di un atto o di uno stato mentale a un determinato oggetto, ovvero la connessione che l'atto o lo stato hanno, in virtù della loro identità, con un certo oggetto, indipendentemente dalla sussistenza di questo eventuale oggetto nella realtà esterna. Ad esempio, se io spero qualcosa, questa speranza o attesa è legata a un oggetto. Il dibattito principale su questo tema si è svolto, di nuovo, tra Searle e Dennett. Il primo sostiene che vi sia una intenzionalità intrinseca (ad esempio quella dei sentimenti umani) e una derivata (ad esempio quella mostrata da oggetti progettati dall'uomo). L'intenzionalità derivata esiste nella misura in cui un agente con intenzionalità intrinseca l'ha progettata o realizzata e, dunque, l'intelligenza artificiale potrà avere soltanto la prima. Dennett ha contestato a questa concezione dell'intenzionalità il fatto di essere pretestuosa e inesistente, ch  l'intenzionalità   una ed unica sebbene funzioni come sistema complesso che pu  essere "scomposto" in sottosistemi via via pi  semplici. Gli stati mentali umani, in questo senso, non sarebbero altro che manifestazioni di fenomeni non visibili o meno visibili perch  di "pi  basso livello" e questa scomposizione potrebbe continuare di fatto fino al limite fisico del singolo neurone. Per Dennett, dunque, la mente non   intenzionale perch  non esiste come ente separato dal corpo: l'intenzionalit  si limita a essere un'espressione verbale e l'intenzionalit  artificiale   concettualmente possibile perch  anche quella che si suppone essere del soggetto   in realt  un fenomeno della materia psichica.

Ci  detto, tuttavia, ancora una volta niente viene precisato in merito alla decisione e a "chi" decide.

4. La sfida della libert .

Un paradosso s'avanza da quanto appena detto: auspicando un dialogo con la filosofia McCarthy chiama in causa la questione del *free will* che, invece, nei filosofi che si interessano di intelligenza artificiale manca perch  essi sono pi  interessati al funzionamento della mente e al determinismo delle azioni. Perci  – ed ecco il paradosso – quando si chiede "chi decide?" si pone una domanda che, nel contesto della riflessione sull'intel-

ligenza artificiale, interessa *meno* la filosofia e *più* altre discipline. Ora, tale interesse interdisciplinare prova che la questione – se posta in modo astratto – non interessa né, dunque, riceve risposta. Si tratta, invece, di una questione che si pone soltanto “in concreto” e che soltanto l’incrociarsi di discipline e metodi diversi chiariscono. Si dovrà dunque concludere che il paradosso di cui si è detto conduce fuori dalla filosofia? No, semplicemente invita a cercare in essa quegli utensili con i quali la filosofia si investe in un dialogo interdisciplinare nel quale, per la natura stessa delle questioni che affronta “in concreto”, non può essere la sola a rispondere, né la sua risposta può bastare. Non si tratta dunque di rinunciare alla filosofia ma di tornare a ciò che di originale essa ha e può dare. E trattandosi della volontà libera e della libertà che intervengono nella decisione, ciò che la filosofia ha da dire e offrire non è poca cosa. Lo mostra un passo della *Critica della ragion pura* di Immanuel Kant che ci permetterà di compiere quel passo in avanti fino a ora mancato.

Nessuno potrebbe contestare lo stretto legame tra la decisione e il soggetto che la definisce “sua” (come nei casi in cui diciamo: “questa è la mia decisione” o “questa decisione è mia”). E poiché decidere è agire, quando si tratta di decisione è fino alla questione “chi agisce?” che occorre retrocedere, ossia fino a quel punto originario che Kant affronta nelle *Antinomie*. Esse riguardano quell’insieme dei fenomeni che chiamiamo “mondo” o “cosmo” (da cui la loro definizione di “antinomie cosmologiche”) e sono tutte costituite da una tesi e un’antitesi irrisolvibili perché mancanti di un’intuizione che le confermi o smentisca. Kant ne enuncia quattro e la terza recita così⁹:

Tesi	Antitesi
La causalità secondo le leggi della natura non è l’unica dalla quale i fenomeni del mondo, nel loro insieme, possano essere derivati. Per la loro spiegazione è necessario ammettere anche una causalità mediante libertà	Non vi è alcuna libertà, ma tutto nel mondo accade secondo leggi della natura

⁹ I. KANT, *Critica della ragion pura*, trad. C. Esposito, Milano, 2012, p. 677.

Nell'un caso (tesi) si afferma che le cause naturali non bastano a spiegare l'insieme dei fenomeni e che occorre introdurre una spontaneità che causa l'insieme dei fenomeni; nel secondo caso invece (antitesi) si afferma che non c'è alcuna causa che sia altra dalla natura e che dunque nessun fenomeno accade a prescindere da cause naturali. Detto altrimenti, il fatto che vi sia o meno una causalità mediante libertà (va precisato che il problema si pone nel contesto della fisica newtoniana che Kant conosceva) non è mostrato da alcun fenomeno "esperibile".

In quanto detto c'è tuttavia un argomento che per Kant è "inaggirabile" e che egli spiega nel modo che rapidamente riassumiamo. Nello spiegare la Tesi, egli precisa innanzitutto che l'idea trascendentale della libertà, ossia l'idea che la causalità secondo natura non sia la sola e che occorra ammettere anche una causalità libera, non coincide con la libertà psicologica del singolo ma è ben più ampia perché riguarda la possibilità di affermare «l'assoluta spontaneità dell'azione»¹⁰. Tale possibilità non è, certo, l'azione che ciascun singolo compie ma è la condizione per cui ciascuno può essere libero. Ora, tale libertà trascendentale va intesa «come il vero e proprio fondamento di imputabilità dell'azione»¹¹ sebbene, così facendo, «essa costituisce la vera e propria pietra dello scandalo di tutta la filosofia, la quale trova difficoltà insuperabili ad ammettere una tale specie di causalità incondizionata»¹².

È dunque "pietra dello scandalo" il fatto che la libertà umana sia possibile soltanto a condizione che una libertà incondizionata si dia; a condizione che si dia, cioè, una libertà prima e assoluta, la quale costituisce la possibilità stessa che l'azione sia imputabile *a me*. Il che equivale a dire che il motivo per cui un'azione "è la mia" non ha la sua condizione di possibilità nell'azione stessa ma in una libertà che, in un certo qual senso, la supera perché è non soltanto libertà del singolo ma è "libertà trascendentale" e tale perché nel mondo si danno eventi che accadono per una spontaneità assoluta. Ancora, tale perché vi sono delle serie di fenomeni che, sebbene una volta che siano iniziate procedono secondo leggi naturali, iniziano per un'originale e previa spontaneità. In definitiva, se non vi fosse questa libertà

¹⁰ *Ibidem*, p. 681, *Osservazioni sulla terza antinomia*.

¹¹ *Ibidem*.

¹² *Ibidem*.

che coinvolge ogni fenomeno pur non essendo essa stessa un “fenomeno manifesto e oggettivo”, un’azione non sarebbe imputabile a me o ad altri.

Ma non è forse questa proprio l’esigenza manifestata da McCarthy quando chiede alla filosofia l’aiuto per affrontare la questione del *free will*? Probabilmente sì, e affrontarla mostra che, qualunque sia il modello (filosofico o ingegneristico) per cui si opta, un punto è inaggirabile: l’azione deve essere imputabile a “qualcuno” e questa è la specificità dell’agire che anche la scienza robotica ricerca. Ora, *decidere* è anche *agire* e tale verbo va assegnato ed esteso a chiunque agisca, si tratti dell’uomo o dell’intelligenza artificiale, dove non a caso gli specialisti della disciplina parlano di “albero di decisioni”. Se tuttavia lo scienziato chiede alla filosofia gli strumenti per affrontare la questione del libero volere, che cosa può chiedere la filosofia all’intelligenza artificiale?

Sicuramente, come è stato detto, le chiede gli strumenti per capire come la conoscenza della mente funziona, e abbiamo visto due diverse vie della *contrée des philosophes* fronteggiarsi adducendo in entrambi i casi valide motivazioni. Motivazioni condivisibili ma irrisolvibili (come le antinomie kantiane) perché aggirano e non dichiarano (più o meno velatamente) la concezione dell’uomo che ciascuno sottende, concezione che è la vera posta in gioco attorno alla quale la filosofia dialoga insieme alle altre discipline, dal diritto e la giurisprudenza all’ingegneria e all’insieme dei saperi che nella complessità attuale convergono.

5. La libertà della decisione.

Rispondere alla domanda di McCarthy è impossibile – con il conseguente incappare nella contraddizione di un’azione senza agente – se non ci si interroga reciprocamente e di nuovo sulla libertà, sulla libera decisione e su un’imputabilità possibile perché *ogni* agire è libero per la libertà stessa del mondo e del cosmo. Ma com’è possibile rimettere in discussione la libertà senza farne un annuncio vuoto come lo era il soggetto trascendentale da cui si è partiti? È possibile se la si concepisce come un *fiat* che ogni volta deve iniziare “davanti a” casi e situazioni sempre diversi. Una libertà che non si concepisse “davanti a” resterebbe un annuncio astratto e proprio questo è il punto manca-

to dalla trasformazione del soggetto occasionato dall'intelligenza artificiale di cui si occupa la *contrée des philosophes*: se per un verso il dibattito sulla descrizione della mente e sul suo funzionamento è ampio, per altro verso in tale dibattito la questione della decisione libera non viene posta ("freedom is not what matters!") perché ciò che conta sono *l'analisi e la descrizione del funzionamento della mente*. *Inaspettatamente*, invece, la questione del libero agire è posta da chi si occupa di intelligenza artificiale.

È posta *qui* perché tale ambito *oggi* non pone problemi di *analisi* ma di *sintesi*, ossia del peso che la raccolta e l'elaborazione di dati *ha* e delle conseguenze che comporta per l'uomo vivente e concreto. Di fronte a questa sintesi "chi decide" non rimane nella neutralità ma prende decisioni di cui non conosce, scientemente o meno, le conseguenze; in ogni caso, egli sceglie, opera e, con ciò, è "chi decide", mostrando fino a che punto la questione sia inaggirabile. Perciò decidere, lo ripetiamo, non accade mai in astratto ma in concreto. Tale atto è il concretizzarsi della libertà che è sempre incarnata pur senza essere ridicibile a fattori fisici; una libertà che è sempre in situazione – storica e culturale nel senso più ampio del termine – senza ridursi ad essa, al punto che ciascuno, descrivendo la libertà e la libera decisione, prende la *propria* decisione.

Certo, la questione concernente "chi decide" e "chi sia il soggetto di imputazione" non ha trovato ancora "una" risposta definitiva e, forse, non la troverà restando, perciò, *aperta*. Ma contro ogni apparenza che condurrebbe a dire che tale apertura è la sua vaghezza, noi sosteniamo che in realtà essa è la sua ricchezza. La sua irrimediabile apertura coincide infatti con l'inaggirabile consapevolezza che essere "davanti a" ogni realtà è la *chance* di "chi decide" e, in tal senso, anche "essere davanti all'intelligenza artificiale" è uno dei modi in cui tale *chance* accade e si concretizza. E decidendosi "davanti a" tale ambito, il soggetto si trasforma non perché si dis-umanizza, come vorrebbero i profeti dell'anti-intelligenza artificiale, ma perché appare chiaramente che 1- esso si concretizza quando decide e che 2- la consapevolezza di sé *si acquisisce e cresce decidendosi* "davanti alle" sfide che gli vengono dalla realtà (di cui anche l'intelligenza artificiale fa parte). *Di inizio in inizio, secondo inizi che non avranno mai fine*.

THE JUDICIAL DECISION BETWEEN LEGAL GAPS AND
TECHNOLOGICAL INNOVATION: SOME SUGGESTIONS
FROM THE 19TH AND 20TH CENTURIES

Monica Stronati

TABLE OF CONTENTS: 1. Artificial intelligence between opportunities and challenges. – 2. The boundaries of liability: from fault-based liability to strict liability and back (to fault-based liability). – 3. Legal certainty, judicial predictability and Justice.

1. Artificial intelligence between opportunities and challenges.

The interesting and profitable meeting with computer engineers has highlighted that artificial intelligence has created enormous benefits in all areas of human life.

However, new technologies complicate the traditional rules for recognising and attributing possible damage caused by intelligent machines. First of all because the same machine construction may involve multiple skills and each contributor may sometimes participate only for a part without knowing the whole machine and its overall functioning. Moreover, artificial intelligence machines have a self-learning capability necessary to act in complex situations. Therefore, the designers of these machines are not able to predict their actions and, in many respects, are not even able to control them.

Artificial intelligence machines have shifted and confused the boundaries of the decision and there is a liability issue if the action causes damage.

The programmer creates the programmes, or algorithms, through which the machines act. Manufacturers concretely develop machines for their production and diffusion. It is difficult for users to prove that a damage results from such technologies working unexpectedly or abnormally rather than

from negligence or recklessness, in order to obtain compensation for faulty product damage. In such cases, it may be difficult to understand who more rationally is to blame: whether the programmer or the producer (or both, if that is the case) or the user himself.

The problem of assuming liability does not exist in the hypothesis of an illegal use of software or algorithms, because in these cases the subject who will be held responsible for the damage is detectable. The question of liability arises when the artificial intelligence machine itself produces damage.

For this reason, the possibility of creating new ad hoc rules for a new category of "algorithm" civil liability is being considered. In 2016, the European Parliament's Legal Affairs Committee (JURI) submitted a legislative proposal to the European Commission asking for the subjective qualification of the most sophisticated intelligent robots as electronic persons. The Commission should explore

«the implications of all possible legal solutions, such as: creating a specific legal status for robots, so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons with specific rights and obligations, including that of making good any damage they may cause, and applying electronic personality to cases where robots make smart autonomous decisions or otherwise interact with third parties independently»¹.

Here, one of the remarkable elements is the need to regulate liability for damage caused by robots and damage caused by algorithms. The European Union has set the objective of addressing civil liability for damage. In February 2017, the European Parliament adopted a resolution with recommendations concerning civil law standards on robotics. With regard to civil liability, the resolution states that:

«Any chosen legal solution applied to the liability of robots and of artificial intelligence in cases other than those of

¹ <https://www.europarl.europa.eu/doceo/document/JURI-PR-582443_EN.pdf?redirect>.

damage to property should in no way restrict the type or the extent of the damages which may be recovered, nor should it limit the forms of compensation which may be offered to the aggrieved party on the sole grounds that damage is caused by a non-human agent.

The future legislative instrument should be based on an in-depth evaluation by the Commission defining whether the strict liability or the risk management approach should be applied. A compulsory insurance scheme, which could be based on the obligation of the producer to take out insurance for the autonomous robots it produces, should be established.

The insurance system should be supplemented by a fund in order to ensure that damages can be compensated for in cases where no insurance cover exists.

Any policy decision on the civil liability rules applicable to robots and artificial intelligence should be taken with due consultation of a European-wide research and development project dedicated to robotics and neuroscience, with scientists and experts able to assess all related risks and consequences»².

Actually, the issue arises at a global level, since the development of artificial intelligence technology is happening at a global level, and the monitoring of possible harmful effects must be performed at that same level. Possible risks no longer have geographical boundaries, so solutions adopted by the EU should be shared at least by the US and China.

The liability system established in the USA differs from the English one, even if they come from the same starting point. In the USA, the tort of negligence is greater and more incisive. Indeed, in the USA, the concept is that the cause of the tort is only an objective violation of behaviour. The injured party must prove the negligence. However, there is an alleged violation of the standard of diligence when the harmful event would not usually occur if behaviour were diligent.

This mechanism benefits the position of the injured party whose burden of proof will be significantly lightened by

² <https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_EN.html#title1>.

simply having to prove that there has been misconduct at regulatory level.

The US liability system differs from all damage compensation systems. Generally, the offender's sentence to compensation aims at restoring the *status quo ante* through its monetary equivalent, while US law also aims to "punish" the offender. Therefore, the damage compensation pursues the double purpose of special and general (exemplary) deterrence.

China, currently engaged in the codification of civil law, in 2010 promulgated the law on "Non-contractual Liability"³. The 2010 law is a well-balanced ensemble of legal elements deriving from different legal systems. The drafting technique is that of European continental law: first, there is a general part where the abstract principles concerning the subject are described. The following part has specific provisions dedicated to the concrete regulation of well-defined liability hypotheses. However, some liability categories have features that are very close to the common law systems. The issue of non-contractual liability in Chinese law is constantly evolving, both because of the interpretations of the Supreme Court and because of the work done by the Chinese civil law scholars.

The solution proposed by the European Parliament is based on strict liability, i.e. the principle of fault is left behind to give priority to the event itself. The solution, therefore, is to create insurance guarantees, also by creating special funds.

From the perspective of the historian of law, the new issues, related to liability for damages, evoke legal matters that were addressed during the 19th and 20th centuries following the industrial revolution.

Basing on historical experience, which I shall explain in the following paragraph, the choice of the strict liability for damage should intervene only in a secondary way with respect to fault-

³ Cf. LEI CHEN - C.H. VAN RHEE (eds.), *Towards a Chinese Civil Code: Comparative and Historical Perspectives*, Leiden-Boston, 2012, especially: YAN ZHU, *The Bases of Liability in Chinese Tort Liability Law – Historical and Comparative Perspectives*, pp. 335 ff., and KEN OLIPHANT, *Uncertain Causes: The Chinese Tort Liability Law in Comparative Perspective*, pp. 395 ff.; GUO, MINGRUI/BI, XIAOXIAO, *The Main Function of Tort Liability Law of the People's Republic of China: To Prevent, to Deter and to Punish Tortious Conduct*, in H. Koziol (Ed.), *The Aims of Tort Law. Chinese and European Perspectives*, Wien, 2017, pp. 43 ff.

based liability. Firstly, because fault-based liability allows the claim of finding possible negligence which could be linked to codes of conduct drawn up by the respective industrial sectors. Secondly, the intervention of the judge would not only be aimed at obtaining compensation, which would be graduated according to the different forms of liability, but above all, it would allow forms of prevention of harmful events by inducing responsible behaviour both in the design and in the quality and safety controls of products.

Such a multilevel solution implies, of course, a virtuous circle between legislator, jurisprudence, doctrine, nevertheless, in the past as well as today, it seems that the ability of the judges is being questioned, as will be stated in the final section of this article.

2. The boundaries of liability: from fault-based liability to strict liability and back (to fault-based liability).

As in the past, blaming programmers or manufacturers for the damage caused by robots means blaming them for the price of technological development⁴.

The issue of liability for damages is, indeed, a central issue in the 19th century, when technological development and industrialisation generated many cases in which determining the fault of a damage was very difficult if not impossible. The 19th century is the “century of machines”, it is a hymn to progress, «la celebrazione della macchina a vapore, della nuova metallurgia, delle strade ferrate capaci di unire i continenti, della navigazione degli oceani rivoluzionata da navi senza vele e da transatlantici propulsi a elica»⁵ (the celebration of the steam engine, of new metallurgy, of the railway roads able to unite continents, of the ocean’s navigation revolutionized by sail-less ships and by propelled transatlantic ships).

Between the 19th and 20th centuries an intense legal debate ignited - over the classic cases of damages caused by railway

⁴ Cf. M. FAIOLI, *Mansioni e macchina intelligente*, Torino, 2018.

⁵ G. CAZZETTA, *Nell’età delle macchine. Artefici, operai, telegrafisti: diritto codificato e incertezze classificatorie dei giuristi*, in *Lavoro e diritto*, 2018, 3, p. 434.

sparks, explosion of steam boilers and accidents at work - which led to a reconsideration of the traditional link of causality between behaviour and responsibility for damage.

England was the first country involved in the industrial revolution, a pioneer in steam rail transport and therefore the first country to approach the problem of burning caused by railway sparks. Technological innovation had created, simultaneously with development, accidents that brought to light the conflict between industrial development and the rights of landowners. Only a simplistic approach can make us think that «the conflict was solved either in favour of railway companies or in favour of land-owners. Conversely, where land property was not so important and the controlling interest was the expansion of the new technology of railway transportation, the conflict would have been resolved in favour of railway companies»⁶.

In England, there was a complication, because the railway companies operated with the permission of the State and they could therefore have a sort of presumption in the correctness of their conduct. On the other hand, the railway companies bought the land at very high prices «and as the railway industry became more competitive, large payments for small amounts of land became more commonplace: the equation of economic development with public interest provided the central ideological justification for the invasion of the land by the railway»⁷.

In England the tort law was created by «the custom of the judges of the superior courts: it developed in a casuistic way, through their case law reasoning»⁸. The main approach, in relation to the damages caused by the spark, had been that of the stringent notion of fault-based liability of railway operators.

Usually the adoption of the principle of strict liability of the

⁶ M. MARTÍN-CASALS, *Technological Change and the Development of Liability for Fault: A General Introduction*, in M. MARTÍN-CASALS (Ed.), *The development of liability in relation to technological change*, Cambridge, 2010, p. 3.

⁷ Ivi, p. 10.

⁸ M. LOBBAN - J. MOSES, *Introduction*, in M. LOBBAN - J. MOSES (Eds.), *The impact of ideas on legal development, Comparative Studies in the Development of the Law of Torts in Europe*, vol. 7, Cambridge, New York, 2012, p. 10.

owner for damage caused by anything on his property is traced back to the case decided by the Exchequer Chamber in 1866 (*Rylands v. Fletcher*). However, there are reasons to doubt the turning point, because «the reason for the strictness in *Rylands v. Fletcher* was that it involved the infringement of property rights: the case would have been argued and decided differently had it involved personal injury»⁹.

Judge Bramwell L.J.'s opinion is interesting because it anticipated «arguments popular a full century later», namely that strict liability had sense if those who created progress also bore the social costs they produced. Indeed, in 1900 a Bill was introduced that «provide[d] for compensation without the need to prove fault, when crops were destroyed by sparks from railway engines». Meanwhile, economists theorised «that a strict liability [was] optimally efficient, as maximizing cost internalization (so that the cash cost of a product/service [represented] the true social cost of its manufacture or provision)»¹⁰.

Another "side effect" of technological progress was the issue of the explosion of steam boilers, which involved mainly workers. It is estimated that between 1865 and 1882, in the United Kingdom, boiler explosions caused 1,051 deaths and 1,519 injuries¹¹. Despite these serious consequences, there is no «mention of exploding boilers in any of the nineteenth-century tort textbooks»¹². Actually, the fact is not surprising because it was very difficult for the workers to bring cases before a court, not only because the legal costs were prohibitive, but also because of a condition of general weakness of the workers in the industrial context.

The serious issue of accidents was addressed by refusing the government's mandatory controls and by opting for self-regulation. As well as in Italy and France, in England associations of steam boiler users were established, which

⁹ J. MORGAN, *Technological Change and the Development of Liability for Fault in England and Wales*, in M. MARTÍN-CASALS (Ed.), *The development of liability in relation to technological change*, cit., p. 46.

¹⁰ Ivi, pp. 47-48.

¹¹ M. MARTÍN-CASALS, *Technological Change and the Development of Liability for Fault: A General Introduction*, cit., p. 12.

¹² J. MORGAN, *Technological Change and the Development of Liability for Fault in England and Wales*, cit., p. 52.

provided inspections aiming at preventing boiler explosions. The Boiler Explosions Act of 1881, intended to «improve investigation into case of accidents, but it did not aim at establishing compulsory boiler inspection»¹³. Compensation came by way of the practice of the "deodands", an «archaic legal doctrine mentioned in the Old Testament, an objective which had caused a human death would be offered up as a sacrifice to God, or rather, in England, forfeit to the Crown». The practice changed into the payment of a sum by the producer who had caused the death. Since the 1830s «In same case [...] juries were deliberately using the deodant to punish the railway, etc. for callous policy decisions. Sometimes, indeed, a jury may have used the threat of a substantial deodand to ensure that an employer properly compensated an accident victim's family»¹⁴. With the Fatal Accidents Act of 1846, the victim's family was allowed to obtain compensation for their loss. The Boiler Explosions Act of 1881 also provided «the awards referring to the amounts corresponding to investigation costs that [...] were to be paid by the negligent boiler owner and which could be reduced when he had already made provision for a dead worker's family»¹⁵. The Boiler Explosion Act cannot be defined as an effective remedy. The Boiler Explosions Act Commissioners enjoyed no real remedial powers, they were empowered to allocate the costs of the inspection but in «awarding costs, the commissioners were guided by moral blame and not formal legal liability, i.e. strict, vicarious liability was not sufficient for a finding of liability under the Act»¹⁶. In practical terms, the principle of strict liability led to more advantageous forms of insurance than the investigations established by law.

The English lawyers placed increasing emphasis on the

¹³ M. MARTÍN-CASALS, *op. ult. cit.*, p. 14.

¹⁴ J. MORGAN, *Technological Change and the Development of Liability for Fault in England and Wales*, in M. MARTÍN-CASALS (Ed.), *The development of liability in relation to technological change*, cit., p. 57.

¹⁵ M. MARTÍN-CASALS, *Technological Change and the Development of Liability for Fault: A General Introduction*, cit., p. 15.

¹⁶ J. MORGAN, *Technological Change and the Development of Liability for Fault in England and Wales*, cit., p. 63.

notion of fault «but they could not ignore the large space left within the common law for strict liabilities»¹⁷.

The Continental Europe experience differed from that of the common law for various reasons, not least for having codified civil liability: «in codified systems, the domain of delictual liability is shared between three actors: the judge, the legislator and the academic developing doctrine. The last word [...] belongs – or seems to belong – to the judge; but the first word is that of the legislator, in deciding whether or not to recognize a delictual liability»¹⁸.

In the legal systems of Continental Europe, the idea of fault-based liability remained central when approaching compensation for victims of damage caused by third parties, however after 1850 continental models diversified.

The Italian art. 1151 c.c. (Civil Code) of 1865 literally translated art. 1382 of the Napoleonic Code: «Tout fait quelconque de l'homme, qui cause à autrui un dommage, oblige celui par la faute duquel il est arrivé, à le réparer» (Any fact of the man who causes damage to others, obliges the one who caused the damage to compensate for it). The article stated the fundamental principle of *neminem laedere*, i.e. an absolute principle of not causing harm to others. It was also a general principle that can be interpreted by the courts in order to protect unjust damage beyond the limitations imposed by law¹⁹.

The necessary elements of fault-based liability were: the damage (*dommage*), the fault (*faute*) and the causal link (*lien de causalité*) between the behaviour and the damage suffered.

Technical development and industrialisation undermined the right to ownership considered "sacred" in the bourgeois systems of the liberal States. The need for strong protection of ownership «fa[ceva] saltare il filtro di equilibrio della responsabilità per colpa che tende[va] ad essere schiacciato

¹⁷ M. LOBBAN - J. MOSES, *Introduction*, cit., p. 11.

¹⁸ J. L. HALPÉRIN, *The process of codification applied to the law of delicts*, in M. LOBBAN - J. MOSES (Eds.), *The impact of ideas on legal development, Comparative Studies in the Development of the Law of Torts in Europe*, cit., p. 168.

¹⁹ G. CAZZETTA, *Responsabilità aquiliana e frammentazione del diritto comune civilistico (1865-1914)*, Milano, 1991, p. 343, available at <<http://www.centropgm.unifi.it/biblioteca/040/volume.pdf>>.

dal riferimento ‘sistematico’ al più vasto ed assorbente principio di ‘non ledere il diritto altrui’, del non soffrire una diminuzione qualunque dei beni senza indennità; principio che si affianca[va] a quello della colpa e lo svuota[va]»²⁰ (blew up the filter of balance of fault liability which tended to be crushed by the ‘systematic’ reference to the broader and more absorbent principle of ‘not violating the rights of others’, of not suffering any reduction in assets without compensation; a principle which went hand-in-hand with that of fault and emptied it). In the case of damage caused by railway sparks, case law applied the principle of *culpa in re ipsa* and, in order to ensure the compensation of damage, created «more a presumption of liability than a presumption of fault»²¹.

Decisions of the courts took into account the ease for railway companies to prove their lack of fault. Sometimes they adopted the strategy of the reversal of the burden of proof, or they presumed fault in the fact that damage occurred, or they presumed negligence in the adoption of preventive measures²².

The judges sometimes recognised that the damage was unavoidable, but the statutory right of inviolability of property (art. 29 and art. 436 c.c. [Italian Civil Code] 1865) had to prevail, thanks to which, indeed, a reasonable indemnity was guaranteed in the case of expropriation for public utility (art. 438 c.c. [Italian Civil Code] 1865).

If the damage could not be avoided by the railway companies, and therefore they were not to blame, it was possible to have recourse to Article 3 of the preliminary provisions of the Civil Code of 1865 «ed elasticizzare al massimo, sino a falsarle, le disposizioni sulla colpa, ma non [erano] tali da piegare la consolidata tradizione interpretativa nella lettura dell’istituto»²³ (and make the provisions on fault as flexible as possible, to the point of distorting them, but these were not capable of breaking the established tradition in

²⁰ Ivi, pp. 346-347.

²¹ C. FAVILLI, *Technological Change and the Development of Liability for Fault in Italy*, in M. MARTÍN-CASALS (Ed.), *The development of liability in relation to technological change*, cit., pp. 193-197.

²² Ivi, pp. 193-194.

²³ G. CAZZETTA, *Responsabilità aquiliana e frammentazione del diritto comune civilistico (1865-1914)*, cit., p. 351.

the interpretation of liability for fault). For instance, the article 574 of the Italian Civil Code, which provided for the prohibition of nuisance towards a neighbour²⁴, was interpreted in an analogical way. The decisions of the courts still tried to find a basis in the principle of fault-based liability, which was re-elaborated through presumptions and, if such a tool revealed itself not being sufficient, recourse was made «all'assolutezza del principio del *neminem laedere* e al tema della ingiustizia di una espropriazione senza indennizzo»²⁵ (to the absoluteness of the principle of *neminem laedere* and to the theme of the injustice of expropriation without compensation).

The courts introduced elements of extra-judicial evaluation into the decisions and constructed the judicial decisions by adapting the codified law to the values and interests shared by society. It was also for this reason that the judges subsequently returned to the rigorous principle of fault-based liability. Since 1915, the orientation of the Roman court of cassation prevailed, which overcame «a formalistic application of article 1151 c.c. and [established] a genuine interpretation of the responsibility for fault. Now the courts found civil liability only if the injured person was able to prove a fact which was imputable to the defendant in a subjective sense (fraud or fault), besides the *iniuria* and the link of causation to the damage»²⁶.

From a merely formal viewpoint, jurisprudential uncertainty and intolerable interpretative abuse by judges should be made known. However, the decision of the judges «had to reflect a number of factors, only one of which was the legal rules, and law should mirror social movements»²⁷; in the second judicial period, the decisions of the courts reflected the predominance of the general interest of the development of transport over the interest of landowners.

²⁴ Cf. C. FAVILLI, *Technological Change and the Development of Liability for Fault in Italy*, cit., p. 198.

²⁵ G. CAZZETTA, *op. ult. cit.*, p. 347.

²⁶ C. FAVILLI, *op. ult. cit.*, p. 199.

²⁷ J. BELL, *Continental European Jurisprudence 1850-2000*, in M. LOBBAN - J. MOSES (Eds.), *The impact of ideas on legal development, Comparative Studies in the Development of the Law of Torts in Europe*, cit., p. 111.

In Italy, the use of steam boilers was lesser than in the rest of Europe, however there was a high number of accidents²⁸. Also in relation to the phenomenon of the explosion of the boilers, solutions and approaches according to the French model were adopted. In addition to the articles of the Italian Civil Code (413 and 574), in 1888 a statute law on Public Safety No. 5888 *decies* was issued, which in Art. 27 provided for a safety certificate for new or restored steam boilers and for a supervision by local governmental authorities and technicians²⁹.

The subsequent regulation of 27th June 1897, No. 890, specified the activities of supervision but, unlike France, did not mention civil liability, so «the reference standard was the general rules on civil liability», or the articles concerning fault-based liability (art. 1151, 1152 and 1153 paragraph 4 of the Italian Civil Code).

The few decisions of the judges on boiler explosion damages followed the approach of the courts in the recognition of liability for damages caused by train sparks. The liability of boiler manufacturers was declared on the basis of the *culpa in re ipsa* also extended to unforeseeable and unavoidable damages. According to the judges, damage did not always involve negligence, «When an error was inescapable or when it was connected to the level of scientific knowledge at the time, these errors should not be equated to fault»³⁰.

If it was true that for legislators of the beginning of the 19th century, the «delictual liability was not perceived to be an important political question», things changed profoundly «after the first effects of the Industrial Revolution»³¹. The interpretative turning-point of the judges was the recourse to the general clause of articles 1382 and following ones of the Napoleonic civil code (articles 1151 and following articles of the Italian civil code), as in the famous *Teffaine* judgment of 16th June 1896. The French Court of Cassation, on the basis

²⁸ Cfr. S. LOLLINI, *Macchina a vapore*, in *Dig. It.*, vol. XV, I, Torino, 1927, pp. 31ff.

²⁹ <https://www.normattiva.it/do/ricerca/semplice/0>.

³⁰ C. FAVILLI, *Technological Change and the Development of Liability for Fault in Italy*, cit., p. 211.

³¹ J.L. HALPÉRIN, *The process of codification applied to the law of delicts*, cit., p. 177.

of the first paragraph of Article 1384, declared the strict liability of the entrepreneur for the explosion of a steam engine. The Court's decision influenced the judgments of the lower courts «but this line of case law was far from constant or universal». For example, in the *Grange* case, of 30th March 1897, the Supreme Court rejected the widow's petition, establishing that the cause of the accident was a fortuitous event³². The new interpretation of the first paragraph of article 1384 aroused an interesting debate in legal doctrine, however «an employee could still be found contributorily negligent, which accounted for a large number of cases. Furthermore, civil procedures were often lengthy and costly, with judges apparently awarding somewhat 'arbitrary' and unpredictable levels of compensation. Ultimately, the influence of article 1384(1) had a brief reign with respect to industrial accidents, as a Law of 9th April 1898 instituted a new regime»³³.

In the German model, the § 823 BGB did not formally provide for any general rule attributing the obligation to pay damages to any person who caused damage to others in an unlawful and culpable manner. In addition to the general elements of illegality and guilt, the German legislator required that a damage occurred to a good expressly protected by law.

The first exception to the principle of fault-based liability dated back to 1838 with the enactment of the PGE (Preußisches Gesetz über die Eisenbahnunternehmungen). The principle of strict liability «was seen as exceptional and even inappropriate for a systematic classification»³⁴.

In Germany, *culpa* was seen as a way of inducing responsible behaviour and, at the same time, a way of encouraging industrial development. The theory of *culpa* was developed by the Pandectist doctrine «with the clear intention to restrict the field of civil liability»³⁵. The *faute* under Art.

³² Y. SALMON, *Technological Change and the Development of Liability for Fault in France*, in M. MARTÍN-CASALS (Ed.), *The development of liability in relation to technological change*, cit., p. 113.

³³ Ivi, p. 114.

³⁴ J.M. SCHERPE, *Technological Change and the Development of Liability for Fault in Germany*, in M. MARTÍN-CASALS (Ed.), *The development of liability in relation to technological change*, cit., p. 134.

³⁵ J.L. HALPÉRIN, *The process of codification applied to the law of delicts*, cit., p. 182.

1382 of the Napoleonic Code was similar to the German *Verschulden*, however the developments were different, because in France through the *faute* liability was extended, on the contrary the «*Verschulden* was the flag used by German jurists to prevent the development of strict liability»³⁶.

In 1871, with the Liability Act of the Reich, the «strict liability for railways, very much along the line of § 25 PGE»³⁷ was accepted.

As is well known, the codification of civil law in Germany took place only in the 20th century. The BGB (Bürgerliches Gesetzbuch) confirmed that responsibility had to involve fault: «Strict liability [...] was a concept alien to the civil law»³⁸. The question of liability was definitively resolved with the transfer «of (private) liability to the public insurance model», through the Accidental Insurance Act of 1884.

Germany was the first to introduce the social insurance scheme, which was later followed by other countries: England with the Worker's Compensation Act of 1897, Italy with the Law of 17th March 1898, No. 80 that established compulsory insurance for industrial injuries, France with the Law, 9th April 1898 concerning liability for industrial accidents and also Spain on 30th January 1900.

The European models of liability were certainly different; however, we can observe a tendency towards a progressive worsening of liability, especially in the historical periods marked by technological changes and industrial development.

Actually, the two models of civil law (the French and the German ones) have been "surpassed" by the compulsory insurance, i.e. by the allegation of liability that did not depend on the causal connection between damage and fault. The critical area of civil liability was represented by the issue of accidents at work. The legal doctrine developed hermeneutical categories that should have guided judges in their decisions (*culpa lievissima*, reversal of the burden of proof, strict liability). As is well known, the solution was taken outside the civil law, by means of a special law. In this way, the judges

³⁶ *Ibidem*.

³⁷ J.M SCHERPE, *op. ult. cit.*, p. 162.

³⁸ *Ivi*, p. 172.

were relieved from case-by-case decisions and went for the insurance solution.

The solution based on strict liability developed the insurance system and automated decisions; it subtracted the governance of new technologies from the legal domain. Renouncing to assess the responsibility of the actors was not without consequences, first of all there was no incentive for a responsible development of new technologies, on the contrary it ended up allowing, *de facto*, a development with very few limits³⁹.

Basically, in the past as today, there is a certain mistrust towards judges and jurisprudence, undoubtedly, writes Cazzetta: «il rapporto con la giurisprudenza appare come il nodo teorico da sciogliere in tema di responsabilità prima e dopo la scelta del sistema»⁴⁰ (the relationship with jurisprudence appears as the theoretical node to be dissolved in terms of liability before and after the choice of the system).

3. Legal certainty, judicial predictability and Justice.

We are late. The era of artificial intelligence is already here, and law must give answers to many open questions, otherwise it will be swept away. The use of artificial intelligence in the law field is indeed being introduced. Thanks to software designed by legal tech start-ups, a large number of documents can be examined very quickly. As a result, lawyers will enjoy easier decision-making processes and relief from repetitive work and low-level tasks, like for instance, providing real-time answers to simple questions on specific topics thanks to chatbots (conversation simulators). It is expectable that law firms will have to acquire competences in the field of innovation and will therefore increasingly acquire engineering competences.

³⁹ A. BERTOLINI - E. PALMERINI, *Regulating robotics: A challenge for Europe*, in EU Parliament, Workshop on «Upcoming issues of EU law» for the IURI Committee, Compilation of in-depth analyses, Publications Office of the EU Parliament, Brussels, 2014 ([https://www.europarl.europa.eu/RegData/etudes/IDAN/2014/509987/IPOL_IDA\(2014\)509987\(ANN01\)_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/IDAN/2014/509987/IPOL_IDA(2014)509987(ANN01)_EN.pdf)).

⁴⁰ G. CAZZETTA, *Responsabilità aquiliana e frammentazione del diritto comune civilistico (1865-1914)*, cit., p. 374.

In Italy as well, we have experimented the “interpretative algorithm”, a mathematical formula invented by the lawyer Luigi Viola⁴¹, that translates article 12 of the preliminary provisions of the Civil Code, that is the judge’s interpretative process in applying the rule to the case, into an algorithm. The aim, it is argued, is not to replace the interpretative activity of the judge, rather to support the jurist and the judges in the decisions, besides reducing the level of litigiousness and pushing towards forms which are alternative to the judicial procedure. This, the computability of the judicial decision, is an ancient aspiration that can be traced back at least to Leibniz, who proposed to «estendere al dominio della giurisdizione l’ideale della calcolabilità universale [...]. Una calcolabilità che comportava piena prevedibilità»⁴² (extend to the jurisdiction domain the ideal of universal computability [...]. A computability that involved full predictability).

The breakthrough changes introduced by digital technology undoubtedly have a heavier impact if compared with other transformations that justice has undergone in the past. For instance, the smart contracts that produce «une sorte de droit robotisé, qui s’applique seul, sans aucune intervention d’une instance juridique» (a kind of robotic law, which applies itself on its own, without any intervention by a legal instance). It represents a logic «qui ne se limite pas à court-circuiter les médiateurs professionnels: elle rêve de supprimer le langage lui-même comme support du droit pour le remplacer par des algorithmes»⁴³ (which is not limited to bypassing professional mediators: it dreams of removing language itself as a support for law and replacing it with algorithms).

We cannot ignore the risks connected to digital technologies applied in the legal field. The questions involved are various and

⁴¹ *Interpretazione della legge con modelli matematici. Processo, a.d.r., giustizia predittiva*, vol. 1, Milano, 2018.

⁴² M. LUCIANI, *La decisione giudiziaria robotica*, in *AIC*, 2018, 3, p. 879, available at https://www.rivistaaic.it/images/rivista/pdf/3_2018_Luciani.pdf; cf. A. CARLEO (a cura di), *Calcolabilità giuridica*, Bologna, 2017, in particular N. IRTI, *Per un dialogo sulla calcolabilità giuridica*, pp. 17 ff. See also, N. IRTI, *Un diritto incalcolabile*, Torino, 2016; A. CARLEO (a cura di), *Il vincolo giudiziale del passato*, Bologna, 2018.

⁴³ J. LASSÈGUE, *Justice digitale. Révolution graphique et rupture anthropologique*, Paris, 2018, p. 146.

very complex. In the extreme forms, predictive justice is a revolution on many levels that may question the human component of the *ius dicere*. Predictive justice and legal tech are born from the penetration of a capitalistic, profit-oriented logic into the legal profession⁴⁴. The ultimate goal of legal tech is to replace legal culture with artificial intelligence.

In my view, we are in a process that originated back in the 18th century, namely the trend to shift the governance of phenomena from the legal domain to that of politics, then to economics and now to technology.

The symptom of this trend is the figure of the judge progressively losing importance, in favour of the political dimension of the decision. Consequently, legal science has also lost its importance in favour of other scientific competences such as statistics, social sciences and hard sciences. These competences no longer support the judge in his decision, but tend to replace him in the decisional process.

There are some similarities with the 18th century also in the rhetorical strategy that emphasises the benefits of the future and devalues the present. Cesare Beccaria, the inspiring father of 19th century codifications, in his *On crimes and punishment* pamphlet exhorts legislators to adopt codes and dismantle the jurisprudential system of *ius commune*, and «let the laws be feared, and the laws only. Fear of the laws is salutary, but fear of men is a fruitful and fatal source of crimes»⁴⁵. The law, of course, must be merely syllogistically applied by the judge. To those who pointed out that "automated" justice would create injustice, Beccaria replied:

«The disorders, that may arise from a rigorous observance of the letter of penal laws, are not to be compared with those produced by the interpretation of them. The first are temporary inconveniences which will oblige the legislator to correct the letter of the law, the want of preciseness

⁴⁴ Cf. A. GARAPON, J. LASSÈGUE, *Digital Justice. Révolution graphique et rupture anthropologique*, Paris, 2018.

⁴⁵ C. BECCARIA, *Essay on crimes and punishments, translated from the Italian; with a commentary, attributed to Monf. De Voltaire, translated from the French*, London, MDCCLXXV⁴, chap. XLI, *Of the Means of preventing Crimes*, pp. 165-166.

and uncertainty of which has occasioned these disorders; and this will put a stop to the fatal liberty of explaining; the source of arbitrary and venal declamations. When the code of laws is once fixed, it should be observed in the literal sense, and nothing more is left to the judge, than to determine, whether an action be, or be not conformable to the written law. When the rule of right which ought to direct the actions of the philosopher, as well as the ignorant, is a matter of controversy, not of fact, the people are slaves to the magistrates. The despotism of this multitude of tyrants, is more insupportable, the less the distance is between the oppressor and the oppressed; more fatal than that of one, for the tyranny of many is not to be shaken off, but by having recourse to that of one alone. It is more cruel, as it meets with more opposition, and the cruelty of a tyrant is not in proportion to his strength, but to the obstacles that oppose him»⁴⁶.

In the same way, today, we are induced to emphasise the goodness of justice elaborated by intelligent machines and algorithms, because it is a democratic justice, since it is foreseeable and optimises results. And on the other hand, it is all too easy to discredit the work of judges, too fallible, or too conditionable, as we read in Beccaria's *Pamphlet*, the judges make the judgments depending on bad or good digestion.

A key point for the supporters of predictive justice is the achievement, at last, of legal certainty, confusing, however, legal certainty with the predictability of judicial decisions or with the necessity, this indeed real, of reducing the time needed for justice to respond⁴⁷. Legal certainty, understood as

⁴⁶ C. BECCARIA, *Essay on crimes and punishments, translated from the Italian; with a commentary, attributed to Monf. De Voltaire, translated from the French, London; chap. V, Of the Interpretation of Laws*, pp. 16-17.

⁴⁷ Cf. *Questione Giustizia trimestrale*, 2018, 4, (<http://questionegiustizia.it/rivista/2018-4.php>), in particular: A. NATALE, *Introduzione. Una giustizia (im)prevedibile?*, pp. 7 ff.; P. GROSSI, *Storicità versus prevedibilità: sui caratteri di un diritto pos-moderno*, pp. 17 ff.; E. VINCENTI, *Massimazione e conoscenza della giurisprudenza nell'era digitale*, pp. 147 ff.; C. CASTELLI - D. PIANA, *Giustizia predittiva. La qualità della giustizia in due tempi*, pp. 153 ff.; C. COSTANZI, *La matematica del processo: oltre le colonne d'Ercole della giustizia penale*, pp. 166 ff.; S. GABORIAU, *Libertà e umanità del giudice: due valori fon-*

predictability, probabilistic though, of the judge's decision, would sacrifice an essential function, that is to decide in the specific case. And it is an essential operation in the separation between the legislator ruling in an abstract and general way and the decision of the judge who applies the general principles to the concrete case.

Even without thinking about Prediction Machines, the mere prediction of the decision, as a support to the judge, can profoundly affect the decision of the judges. It is naïve to think that a judge cannot be conditioned by the expectations of the parties, who are well informed by the algorithms. The judges could find some relief if they could attribute the weight of uncomfortable decisions, for example precautionary measures, to machines, intelligent but nevertheless still conditionable, fallible and hardly liable.

In my opinion, a possibility opens up here to give back a key role to the interpreter that cannot be replaced or conditioned by any automatism. Jurists are today called to guarantee not the certainty of automatic application of the law, but the certainty of the correct decision in the concrete case.

Judges should enhance, with their legal cultural background, mechanisms and techniques already present in the legal system, and should understand the new application. Naturally, this calls into question the legal education of jurists. University education must be renewed and it must give priority to interdisciplinary and multidisciplinary education.

True innovation is a new juridical Humanism that makes men the key to the civilisation of trials, justice and society, also thanks to the aid of artificial intelligence. «Se devo scegliere qualcuno di cui *non* fidarmi» (If I have to choose somebody I cannot trust), writes Luciani, «personalmente, scelgo il giudice. E scelgo che sia un essere umano. Umano e consapevole dell'importanza, certo, ma anche dei limiti, della sua funzione»⁴⁸ (personally, I choose the judge. And I choose that it's a human being. Human and conscious of the importance, certainly, but also of the limits, of his/her function).

damentali della giustizia. La giustizia digitale può garantire nel tempo la fedeltà a questi valori?, pp. 200 ff.

⁴⁸ M. LUCIANI, *La decisione giudiziaria robotica*, cit., p. 893.

LÉGIFÉRER SUR L'INCONNU. À PROPOS DES SYSTEMES D'ARMES LÉTAUX AUTONOMES

Nicolas Haupais

SOMMAIRE: 1. Introduction. – 2. Que faire ? Comment légiférer ? – 3. Les aspects juridiques de l'interdiction/limitation de l'emploi. – 4. La dynamique de l'interdiction: réciprocité et obligations interdépendantes.

1. Introduction.

L'humanité est-elle à la veille d'un bouleversement sans précédent dans les affaires militaires ? Tel est l'enjeu, pour certains, de l'automatisation de plus en plus poussée des systèmes d'armes. Elle suscite autant d'espoirs que de méfiance. Mais il convient de garder la tête froide. Et l'on ne sait guère si ce sujet est le plus énorme dans le cadre du droit international des armements, qui bouleversera toutes les acquis et certitudes ou si l'arrivée de la robotisation n'aura d'effets que marginaux ou parfaitement gérables en l'état du droit existant. Sans compter qu'il est tout-à-fait envisageable qu'elle ait des effets totalement inattendus et en aucun cas anticipables... La difficulté à saisir une réalité à venir est d'autant plus grande que la robotisation véhicule un imaginaire particulièrement développé et souvent effrayant. Tout se mélange, les drones, les cyborgs, les robots à forme humanoïde ou proches de l'animal. Ces confusions sont très gênantes. Il n'est pas possible de traiter de la même manière la question du déportement du pilote (un drone a un pilote humain mais il n'est pas embarqué), des systèmes semi-autonomes, autonomes au sens des Etats-Unis, les robots en tant que tels etc... Quoi qu'il en soit, la robotisation, l'automatisation – concept particulièrement difficile à définir au demeurant et renvoyant à des réalités extrêmement

différentes – existent déjà et s'insinuent dans les systèmes militaires. Mais, parce qu'on considère qu'elles pourraient véritablement exploser, en raison en particulier des progrès de l'intelligence artificielle, se pose la question d'un point qu'elles ne devraient pas franchir, tant elle serait source de dangers.

L'autonomie suscite un emballement éditorial conséquent, et dans tous les domaines¹. Juridiquement une grande activité a suivi la promesse (ou le cauchemar) de telles armes, au sein du Forum de San Remo et du CICR notamment. Au sein des Nations unies, la question a été portée devant la Conférence du désarmement dans le cadre de la Convention sur les armes classiques. Pour l'instant, nous n'en sommes qu'au moment des débats. Aucune régulation ne semble à court terme envisagée et l'on est loin de la convention d'interdiction que certains appellent de leurs vœux et dont la portée serait de toute manière forcément limitée.

L'automatisation saisit le droit ou plutôt le perturbe. A ce titre, elle suscite un certain nombre d'inquiétudes et des interrogations récurrentes. Elles concernent en particulier, formulées en termes juridiques, la question du respect du droit international humanitaire. Elles concernent également une question centrale dans tout système juridique, celle de la responsabilité. Globalement, émerge l'idée que l'introduction d'une entité non humaine dans le processus de désignation d'une cible et de traitement de celle-ci éloigne les éventuels responsables des actes dont ils auraient à répondre². L'objet de cette brève contribution consiste à identifier l'originalité de l'approche de l'automatisation dans le champ du droit international. En effet, deux contraintes particulières pèsent

¹ V. en particulier V. BOULANIN - M. VERBRUGGEN, *Mapping the Development of Autonomy in Weapon Systems*, SIPRI, 2017; Y. DE FRANCE, *Le droit à l'épreuve de l'autonomisation des systèmes militaires robotisés terrestres*, Thèse (dact.), Université de Poitiers, 2019.

² V. sur ce point: J. ANCELIN, *Les systèmes d'armes létaux autonomes (SALA): Enjeux juridiques de l'émergence d'un moyen de combat déshumanisé*, dans *Revue des droits de l'homme*, 2016, Actualités Droits-Libertés 2016; E. POMES, *Les systèmes d'armement létaux autonomes et le droit international humanitaire*, dans *AFRI*, 2016, pp. 243-258; A. CAROLLO, *L'engagement pour des actes imputables aux «robots tueurs»: questions choisies*, dans *AFRI*, 2016, pp. 259-266.

dans les travaux relatifs à l'autonomisation. La première est l'énorme incertitude dans la détermination des réelles virtualités de l'autonomie. La réglementation est étroitement dépendante des capacités de la machine et des systèmes automatisés. Elle l'est aussi de la volonté des acteurs eux-mêmes, et en particulier de militaires, qui ne sont peut-être pas prêts à céder (toute) la place à des entités non-humaines. Donc se pose la question de l'identification du mal que l'on veut éviter, de l'existence réelle de ce mal et des moyens de le conjurer. La seconde contrainte est l'équilibre qu'il convient de trouver entre des impératifs apparemment contradictoires. Cette recherche de l'équilibre est tout-à-fait banale dans le droit; la conciliation entre des impératifs sociaux opposés n'est pas sa raison d'être mais elle constitue sa manière d'être. Se pose seulement la question, sur l'autonomie des systèmes militaires, de l'identification des intérêts sociaux en cause. Se pose surtout la question de savoir si certains d'entre eux ne seront pas jugés à ce point essentiels qu'ils empêcheront toute régulation véritable. La réglementation des armements appelle à un arbitrage entre des préoccupations dont les plus évidentes sont les aspects moraux de leur utilisation et la détermination de ce que les Etats perçoivent comme nécessaires à leur sécurité. Des Etats ont ainsi pu accepter de renoncer à des armes à sous-munitions ou aux mines anti-personnel et refuser des engagements contraignants sur l'arme nucléaire... Les arbitrages sont donc délicats, sachant qu'un Etat fera souvent prévaloir une logique de préservation de sa sécurité. Et il ne faut pas ignorer le fait qu'en matière d'armement, les comportements sont profondément réciproques. Les Etats n'agissent pas isolément; ils se regardent. Et ce que l'un fera aura des répercussions sur les positions de tous les autres. Ce sont ces contraintes, ces difficultés qu'il conviendra d'explorer dans le présent article.

Nous verrons que les questions de définition sont absolument centrales dans la question de la régulation. Nous y reviendrons donc. Pour les partisans de cette dernière, l'automatisation est problématique, notamment lorsqu'elle deviendrait totale, incluant alors «la décision» de tuer. L'on conçoit que toute automatisation n'est pas condamnable dans les choses militaires. Mais, juridiquement, l'autonomie sera la part maudite. Il conviendra alors de fixer un cap au-delà duquel on verserait dans l'inacceptable. Pour approcher

l'autonomie, partons des définitions développées par les Etats-Unis. Le Département de la défense a développé son concept d'autonomie dans une directive du 21 novembre 2012. Ce texte distingue deux types de degré dans l'autonomie et la robotisation, la semi-automatisation et l'automatisation proprement dite. Le texte pertinent mérite qu'on le cite intégralement. Le système autonome est défini ainsi:

“Autonomous weapon system. A weapon system that, once activated, can select and engage targets without further intervention by a human operator. This includes human-supervised autonomous weapon systems that are designed to allow human operators to override operation of the weapon system, but can select and engage targets without further human input after activation.”

Quant au système semi-autonome, il se caractérise par l'élément suivant:

“A weapon system that, once activated, is intended to only engage individual targets or specific target groups that have been selected by a human operator. This includes: Semi-autonomous weapon systems that employ autonomy for engagement-related functions including, but not limited to acquiring, tracking, and identifying potential targets; cueing potential targets to human operators; prioritizing selected targets; timing of when to fire; or providing terminal guidance to home in on selected targets, provided that human control is retained over the decision to select individual targets and specific target groups for engagement; “Fire and forget” or lock on-after-launch homing munitions that rely on TTP to maximize the probability that the only targets within the seeker's acquisition basket when the seeker activates are those individual targets or specific target groups that have been selected by a human operator”.

Dans le cadre de ces armes semi-autonomes, il s'agit à la fois d'assistance du soldat dans l'identification et la désignation des cibles ou de l'automatisation de la phase finale de l'action, sachant que les choix initiaux auront été réalisés par des êtres humains. Notons seulement que du côté

des Etats-Unis, l'engagement dans l'automatisation, couplé à la confiance dans l'intelligence artificielle, est assumé comme un objectif prioritaire depuis l'avènement du «Third Offset Strategy» de 2014. Il s'agit de considérer que les Etats-Unis doivent maintenir un avantage technologique qui doit garantir leur supériorité stratégique, encore que le premier Offset qui date, grossièrement, de la présidence Eisenhower, renvoyait à la nécessité de compenser le déficit de puissance sur le plan conventionnel en misant massivement sur du nucléaire très performant, et non à la nécessité de maintenir un avantage mais de rattraper un retard. La question n'est donc pas de déterminer si l'intelligence artificielle doit faire son entrée dans les affaires militaires mais de déterminer à quelles limites elle sera soumise.

2. Que faire ? Comment légiférer ?

Quand on évoque une interdiction préventive ou des limitations d'emploi, se posent aux différents acteurs (et en particulier aux Etats) deux questions: faut-il interdire (donc faut-il craindre) ? Que faut-il craindre (tout ne pose pas des problèmes éthiques, moraux équivalents) ? Cela n'est pas toujours clairement formulé, mais ce qui pose évidemment le plus de problèmes, c'est le robot-tueur, la machine qui pourrait prendre, de manière autonome, la décision de tuer, dans laquelle l'homme est «out of the loop», hors de la boucle.

Il convient de noter l'originalité de la démarche d'appréhension de l'autonomisation car deux éléments, souvent absents, se combinent. Tout d'abord, les interdictions ou limitations d'usage des armes sont souvent réactives et non préventives. L'emploi d'armes chimiques est interdit en 1925 après leur utilisation pendant la première guerre mondiale. Les armes incendiaires voient leur emploi limité en 1980 après leur large utilisation au Vietnam. Il en est de même pour les mines anti-personnel et les armes à sous-munitions: c'est l'expérience de leur emploi qui montre leur caractère horrible. Certes, toutes les prohibitions, que ce soit de possession, de stockage ou d'emploi ne sont pas nécessairement réactives. Les maladies ont été utilisées comme armes de guerre mais l'arme bactériologique pas vraiment en tant que telle. Les balles explosives ont été utilisées mais pas les armes dont les

éclats ne sont pas localisables aux rayons X. Mais, globalement, l'interdiction suit une première utilisation. Ensuite, les armes interdites préventivement ont des effets connus et parfaitement quantifiables. Ainsi, les armes à laser aveuglantes sont interdites alors qu'elles n'ont jamais déployées. Mais leur effet est parfaitement connu et anticipé. Et parce qu'elles sont perçues comme des armes horribles, elles sont interdites. Dans le cas des armes autonomes, personne ne sait vraiment où elles mènent. De nombreuses incertitudes pèsent sur elles. Jusqu'où ira l'intelligence artificielle ? Ne sera-t-elle finalement pas décevante ? Un autre élément apparaît à ce stade: les Etats veulent-ils d'une véritable robotisation qui irait jusqu'à écarter l'homme ? Bref, c'est sans doute l'une des premières fois que l'on envisage sérieusement d'interdire une arme dont on n'a guère d'idées concrètes.

Les questions se compliquent encore lorsqu'il s'agit de définir l'objet même que l'on prétend saisir. Dans une session organisée par la conférence sur les armes classiques³ et consacrée à l'élaboration d'une définition de travail ('working definition'), un certain nombre d'experts interviennent. La variété de leurs approches frappe. L'un verra l'autonomie comme la capacité à s'auto-diriger, l'autre comme dépendant de la prévisibilité d'un système, d'autres encore évoquent la notion de contrôle humain, de jugement humain. Il s'agit dans ces derniers cas de faire dépendre l'évaluation de l'autonomie du degré de contrôle que l'homme peut exercer sur la machine. Les débats entre les délégations étatiques ne dissipent pas le malaise définitionnel, d'autant que les arrière-pensées influencent fortement les approches. Pour certains, l'autonomie n'est pas une difficulté en soi et il faut alors se contenter d'encadrer son développement pour éviter des abus. Pour d'autres, l'autonomie doit être combattue. Cela pèse inévitablement sur le travail de définition. Notons seulement qu'il est considérablement entravé par les incertitudes qui pèsent sur ces nouveaux systèmes.

Le sujet est donc très fuyant parce que les incertitudes entravent toute capacité de régulation. Si l'on veut interdire, il

³ V. Report of the 2016 Informal Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS), Submitted by the Chairperson of the Informal Meeting of Experts.

faut identifier un mal à éviter et des objets qui pourraient le commettre. Il s'avère difficile de construire un argumentaire avec tant d'incertitudes, alors que l'automatisation est à la fois relative dans de nombreux cas et qu'elle existe déjà dans des systèmes armés. Elle n'est pas un absolu. Elle est un *continuum* et peut renvoyer à des réalités extrêmement diverses⁴. Y. de France identifie ainsi dix stades dans l'autonomisation. Se pose donc la question de la détermination d'un seuil au-delà duquel on verse dans l'inacceptable et donc dans la prohibition. Mais ce stade est peu aisé à déterminer, d'autant que la légitimité morale d'une arme est également contextuelle et dépend de multiples facteurs. Quelques exemples l'illustrent. Notons qu'il s'agit de systèmes existants et déployés. Mais les problèmes éthiques pourraient être les mêmes avec des systèmes pleinement autonomes.

Ainsi, les systèmes anti-missiles embarqués à bord des navires ou les systèmes modernes des contre-mesures visant à protéger des véhicules blindés ou des chars, ce qu'on appelle les APS (Active Protection System) sont des systèmes pleinement autonomes. Ils sont jugés légitimes pour deux raisons. Tout d'abord, ils sont activés surtout pour contrer non des hommes mais des objets (missiles, torpilles, drones suicide): ils ne sont pas tueurs. Par ailleurs, ils corrigent une déficience de l'homme, ou en tout cas un différentiel de capacité qui permet à la machine de remplir une mission dont il est incapable du fait de sa trop grande lenteur. Or, une attaque menée par un missile doit être déjouée dans un délai très bref. Donc la contrainte technique décide de la légitimité

⁴ V. sur ce point: Y. DE LA PANOUSE, *Le droit à l'épreuve de l'autonomisation des systèmes robotisés terrestres*, Thèse (dact.), Université de Poitiers, 2019. L'auteur distingue dix niveaux dans l'autonomisation de la sélection des décisions et des actions: «1. L'ordinateur n'apporte aucune assistance et l'être humain est responsable de toutes les décisions et actions, 2. L'ordinateur propose un ensemble complet d'alternatives pour les décisions et actions, 3. Il réduit la sélection à quelques alternatives ou 4. Il suggère une alternative, 5. Il exécute cette suggestion si l'être humain la valide ou 6. Il laisse un temps limité à l'être humain pour stopper l'exécution avant toute exécution automatique, 7. Il exécute automatiquement et ensuite informe obligatoirement l'être humain, 8. Il informe l'être humain seulement si cela lui est demandé ou 9 Il informe l'être humain si seulement il (l'ordinateur le décide), 10 L'ordinateur décide de tout, agit de façon autonome et ignore l'être humain».

morale. Notons que des effets létaux ne peuvent être exclus puisqu'ils pourraient être actionnés en cas d'attaques suicides avec homme embarqués d'ailleurs. Dès lors que l'homme peut réagir parce qu'il en a le temps, il prend la décision. C'est le cas des systèmes de système de déni d'accès, anti-missiles, tels «Iron Dome» ou «Patriot». On parle de systèmes semi-autonomes. Par ailleurs, un système automatisé de détection de la présence humaine dans la zone démilitarisée entre les Corée du nord et du sud a été mis en place. Une double justification a été apportée sur le plan éthique. D'une part, la présence d'un humain dans une telle zone est forcément une anomalie et il ne peut pas s'agir d'un civil. D'autre part, la machine identifie et signale ce qui lui apparaît correspondre à une personne humaine mais elle ne prend pas la décision de tuer: l'homme «in the loop», dans la boucle. Il faut noter que des projets d'armes fureteuses («*loitering weapons*») ont été développés, mais sur la question de l'autonomie, les recherches ont été arrêtées à la demande des militaires qui n'en veulent pas et qui veulent conserver de l'humain.

Evidemment, il serait possible de considérer que l'appréciation ne peut être transposée à ce qui ferait l'objet d'une interdiction, à savoir le système pleinement autonome. Après tout, dans ces systèmes, il y a en quelque sorte et nécessairement une décision préemptée de donner la capacité à la machine de réaliser une action létale. Si cette décision préemptée est réellement décisive et que la machine a les capacités techniques de la réaliser, pourquoi considérer qu'elle est par principe illégitime sur le plan éthique ?

Toute automatisation n'est donc pas condamnable. Ce qui pose problème, ce serait donc des systèmes entièrement automatisés, la création d'un cyborg autonome, de véritables soldats qui interviendraient le cas échéant mêlés à des êtres humains, des systèmes d'essaims de drones autonomes qui saturent des systèmes de défense, attaquent en masse, réalisent des missions suicide.

L'automatisation fait quoi qu'il en soit l'objet de critiques extrêmement fortes, d'autant qu'elle véhicule des imaginaires extrêmement forts. La mobilisation est importante sur cette thématique, en particulier celle de l'ONG «Stop Killer Robot». Des Etats sont également très en pointe dans la volonté d'interdiction. On rappellera également les appels nombreux contre ces nouveaux systèmes, de scientifiques,

comme S. Hawking, ou d'industriels, tel Elon Musk. Dans une lettre de 2017, une centaine d'industriels, de chercheurs, de scientifiques ont manifesté leur grande inquiétude à l'égard du développement de l'autonomie. Ils soulignent:

“Lethal autonomous weapons threaten to become the third revolution in warfare. Once developed, they will permit armed conflict to be fought at a scale greater than ever, and at timescales faster than humans can comprehend. These can be weapons of terror, weapons that despots and terrorists use against innocent populations, and weapons hacked to behave in undesirable ways. We do not have long to act. Once this Pandora's box is opened, it will be hard to close. We therefore implore the High Contracting Parties to find a way to protect us all from these dangers.”

On le voit, l'approche est ici globale et traduit une réflexion non sur un droit de l'homme mais sur la notion de protection de l'humanité⁵. Ici, le droit de l'armement rejoint des préoccupations que l'on retrouve par exemple dans les normes relatives au réchauffement climatique: il s'agit de conjurer un risque que fait peser sur elle la technologie, des évolutions liées au progrès technique sur l'humanité. Ces discours sont ceux de Hawking ou de Musk qui évoquent des risques liés à la question de l'évacuation de l'humain des processus décisionnels, un risque de perte de contrôle, la création d'un concurrent que l'on aurait en plus immédiatement doté de compétences militaires. Il s'agit de penser les risques de dissémination et de développement de capacités au profit de groupes terroristes et extrémistes. Le risque ultime serait de créer une nouvelle espèce qui pourrait finir par supplanter l'homme.

⁵ Rappelons également qu'en juin 2018, Google, sous la pression de ses employés, a renoncé à renouveler des partenariats avec des industries de défense au nom des dangers liés à l'AI dans le domaine militaire. V. la lettre d'employés de Google adressée à son dirigeant [<https://static01.nyt.com/files/2018/technology/googleletter.pdf>]. Il s'agissait de la participation de la société à un projet dit «Maven» qui permettait d'améliorer la détection de véhicules pour les drones. Il ne s'agissait nullement, d'ailleurs, de systèmes d'automatisation de la décision d'engager un tir.

L'argumentation, finalement, relève plus d'une approche *jus ad bellum* que *jus in bello*. Sur ce dernier point, les critiques sont également présentes et peuvent se présenter ainsi. Certains voient une impossibilité morale, éthique de donner une capacité que l'on peine à qualifier en vocabulaire juridique: droit ? pouvoir ? prérogative ? faculté de tuer un être humain ? Il existerait une forme de droit de ne pas être tué par un robot. C'est sans doute l'approche du Parlement européen dans une Résolution du 12 septembre 2018 sur les systèmes d'armes autonomes⁶ qui évoque «qu'une décision à caractère létal ne peut se prendre sans l'intervention et le contrôle d'un être humain, étant donné que ce sont les humains qui restent comptables de toute décision concernant la vie ou la mort», et que «les machines et les robots ne sont pas en mesure, contrairement aux humains, de prendre des décisions fondées sur les principes légaux de distinction, de proportionnalité et de précaution».

D'autres ont une approche pragmatique et développent une approche comparative de l'homme et de la machine. Dans ce cadre d'analyse, les armes autonomes ne doivent pas être déployées et autorisées parce qu'elles seraient nécessairement inférieures à l'être humain sur certains plans. Il s'agit là d'un discours à double niveau. Le premier est technique et renvoie à l'incapacité de la machine à identifier certaines subtilités aisément perceptibles pour l'homme: qu'est-ce qu'un combattant ? Comment une machine peut interpréter des signes ? Attacher une signification à des comportements ? Le second, certes très proche du premier, inclut la dimension éthique dans le rejet de l'automatisation et de la robotisation. La machine est perçue comme incapable de ressentir certains affects, sensations, sentiments qui influent positivement sur le comportement du combattant, tels que la pitié, l'indulgence, la capacité à trouver des biais pour réaliser un objectif sans passer par des décisions fatales et peut-être, dans certaines situations critiques, à négocier.

Ce discours éthico-technique est évidemment convaincant. Mais il ne résout évidemment pas toutes les difficultés, ce qui est d'ailleurs un trait commun à tous les discours qui traitent,

⁶ 2018(2752)(RSP))

non pas d'une arme existante, mais de systèmes dont la réalité est encore extrêmement fuyante. Et sur ce plan, l'autonomisation pose des difficultés difficilement solubles. En ce sens, il est difficile de tenir un discours ferme parce que manque un référent stable, celui de la réalité des capacités de la machine et des intentions réelles de ceux qui les développent ou qui les mettront en service opérationnel. Le discours d'opposition est donc confronté à un discours de légitimation. Il ne faut toutefois pas nécessairement opposer trop frontalement les deux approches. Evidemment, si l'on raisonne en termes binaires, l'un emportera sur l'autre; l'automatisation sera mise en place dans la décision de tuer ou pas. Mais le discours d'opposition à l'autonomisation peut être intégré dans la décision d'y recourir. En d'autres termes, il peut avoir un effet modérateur. Et il sera intégré par les partisans qui comprennent qu'ils doivent tenir compte d'un certain nombre d'impératifs. Même perdant, ce discours ne sera pas inutile. Du côté des partisans de l'automatisation, les arguments peuvent se présenter ainsi.

La première approche renvoie à la réversibilité éthique des affects liés à la qualité d'homme. Certes, les êtres humains sont capables de pitié, d'empathie, de sentiments tels que l'honneur. Mais, inversement, ils sont sujets à des sentiments qui peuvent amener à des erreurs d'appréciation ou à des actions intentionnellement nuisibles par fatigue, peur, lassitude. Il existe ainsi une violence dite autopoïétique, c'est-à-dire injustifiée au regard des nécessités de la guerre mais dont le ressort est le plaisir du soldat qui s'y livre⁷. La qualité de combattant inverse le rapport à la violence. Normalement interdite, elle est autorisée, quoique contenue, par le droit humanitaire dont on sait qu'il est fondé sur le principe d'humanité. D'où des phénomènes de sadisme, de gout à la violence, de plaisir dans le fait d'infliger la souffrance, la perturbation dans le jugement provoqué par la vengeance, la colère, dont on peut estimer que les machines seraient exemptes. Pour certains, la machine peut faire aussi bien que l'homme, voire mieux dans certains domaines. A ce titre, la question de la morale n'est qu'une question technique. S'il est

⁷ V. en particulier: S. NEITZEL, H. WELZER, *Soldats. Combattre, tuer, mourir: Procès-verbaux de récits de soldats allemands*, 2013, Paris.

possible de reproduire un comportement qui traduit une conscience morale élevée par une machine, la question morale sera résolue, peu important alors que la machine n'ait pas de sentiments moraux. Certains considèrent que l'arme autonome est pleinement admissible dès lors que la machine est dotée de capacités éthiques. C'est l'approche, en particulier, de Brice Erbland: «Un SALA ne pourra jamais posséder une morale dans le sens où il ne sera pas une émanation de sa conscience. Mais il pourra faire en sorte de la simuler, et l'important reste que son comportement soit le même que celui d'un homme moralement droit, ou plutôt qu'il soit au moins aussi bon». Un robot est admissible quand il est enté d'une «éthique artificielle»⁸. Et c'est alors que l'on peut penser non une infériorité de la machine, mais une véritable supériorité qui pourrait l'ériger en œuvre morale. En effet, si la machine peut agir moralement, se révèlent alors des avantages à son profit, libérée qu'elle est des forces obscures qui peuvent pousser un homme à agir contre la morale ou des principes juridiques. Si elle agit moralement, elle agit toujours moralement. Et l'on sait qu'au combat, le respect scrupuleux du droit humanitaire passe aussi par la capacité du combattant à prendre des décisions judicieuses. Il doit faire usage de son intelligence, elle-même arrimée à la mobilisation de ses sens. La technique corrige sinon les infirmités de l'Homme, au moins ses limites. Cette confiance dans la technique peut alors prendre deux formes. Elle peut se traduire par une volonté d'accroître les capacités de l'homme, c'est ce que l'on appelle le «soldat augmenté». Elle peut se traduire par la création d'une autonomie de la machine qui remplace alors l'être humain. On pourrait d'ailleurs remarquer que la formation militaire vise à une certaine robotisation, automatisation de l'humain: acquisitions de réflexes en situation de stress intense et de désorientation sensorielle, maîtrise de ses instincts, capacité à n'infliger que des dommages licites liés dans le seul but de réaliser la mission. Finalement puisque les combattants sont déjà un peu «robotisés», pourquoi ne pas s'arrêter en si bon chemin ? L'OTAN, notons-le, rentre dans cette logique et considère que la question de l'automatisation est avant tout un

⁸ B. ERBLAND, *Les robots tueurs. Que seront les soldats de demain ?*, 2018, Paris.

problème technique. Dès lors qu'elle permet le respect du droit international humanitaire, elle ne pose pas de problèmes fondamentaux⁹.

Cette approche doit être différenciée d'une conception qui vise à minimiser le problème non sur le plan éthique mais sur le plan technique. Il s'agit de considérer que la robotisation n'est pas dangereuse en raison de l'absence de volonté qu'elle le soit. Et elle s'appuie sur un concept, celui de la «fonction automatisée», distinguée du «système autonome» en tant que tel. Cette distinction est réalisée en particulier aux Etats-Unis et a été reprise dans l'approche de l'OTAN. Robert Work, personnalité particulièrement impliquée dans le «troisième Offset», note également¹⁰:

“When people hear me talk about this [autonomy], they immediately start to think of Skynet and Terminator, I think more in terms of Iron Man ... A machine to assist a human, where a human is still in control in all matters, but the machine makes the human much more powerful and more capable. According to this narrative, the purpose of pursuing advances in autonomy is not to remove humans from unmanned systems altogether, but to change their role by creating new models of human-machine collaborations in which the capabilities of both humans and machines can more effectively complement each other.”

Ce discours consiste donc à affirmer que les militaires ne veulent pas d'une automatisation qui irait trop loin dans le processus de substitution de leurs propres décisions par des décisions qui seraient prises par des machines. Il est omniprésent, en particulier dans l'approche américaine. Karl Chang, du Département d'Etat a ainsi précisé la position américaine.

⁹ Headquarters Supreme Allied Commander Transformation, A. P. WILLIAMS, P. D. SCHARRE (eds.), *Autonomous Systems Issues for Defence Policymakers*, p. 5: “Humans and machines excel at different cognitive tasks, and the best model will invariably be a blend of the two”.

¹⁰ Cité in V. Boulanin, M. Verbruggen, *op. cit.*, p. 64.

“As we explain in our working paper, we do not think that these assumptions are accurate. As a factual matter, we think that the use of autonomy in weapon systems can improve the degree of control over the use of force. Moreover, whether a decision is “delegated” in some sense to a machine has more to do with how the weapon is used rather than whether the weapon system itself has an autonomous function. A weapon with a function for selecting and engaging targets can be used without delegating any decision-making to the machine, in the sense of substituting the human’s decision with the machine’s decision.”¹¹

Les Etats-Unis développent une approche qui se présente comme concrète et pragmatique. Reprenons ici le fil des propos de K. Chang:

“We recommend that States focus our discussions on ensuring that weapon systems with autonomous functions operate as intended under appropriate levels of human judgment, reducing the risks of unintended engagements or accidents. This is a clear and practical objective that can have real-world benefits by, for example, facilitating the exchange of best practices in autonomy and weapon system safety.

Our second point is that we do not think that human or “manual” control is the correct issue to focus on. The concept of “human control” is subject to divergent interpretations that can hinder meaningful discussion.

For example, existing international humanitarian law instruments, such as the CCW and its Protocols, do not seek to enhance “human control” as such. Rather, these instruments seek, *inter alia*, to ensure the use of weapons consistent with the fundamental principles of distinction and proportionality, and the obligation to take feasible

¹¹ U.S. Delegation Statement on Human-Machine Interaction at the Meeting of the Group of Governmental Experts of the High Contracting Parties to the CCW on Lethal Autonomous Weapons Systems Delivered by Karl Chang Associate General Counsel, U.S. Department of Defense, Geneva, August 28, 2018.

precautions for the protection of the civilian population. Although control over weapon systems can be a useful means in implementing these principles, “control” is not, and should not be, an end in itself.”¹²

On conçoit qu'un tel discours constitue une forte légitimation. Il consiste à banaliser l'automatisation en conjurant le spectre d'une machine qui serait autonome: il n'existe pas de système qui soit en tant que tel autonome; il n'existe que des fonctions autonomes au sein d'un système dans lequel l'homme est intégré, de manière variable en fonction des circonstances. L'insistance sur l'interface homme/machine, le premier n'étant jamais exclu mais seulement connecté au second, présent pour l'assister, le soulager, le rendre plus clairvoyant et efficace, permet également de rendre vaine une véritable possibilité d'interdiction préventive. En effet elle tend à rendre contextuelle l'approche de l'automatisation. Elle vise à considérer que, dans certains cas, l'autonomie est dangereuse mais que, dans d'autres, elle ne le sera pas. Or, une norme juridique, surtout préventive, vise à conjurer un mal, un danger, que l'on juge présents dans des conditions normales et courantes d'utilisation. L'approche contextuelle permet de ne pas disqualifier en soi un type d'équipement qui n'a pas vocation en soi à être interdit. Ce qui pourra seulement alors être envisagé, c'est l'interdiction d'automatiser certaines fonctions ou la limitation d'emploi de systèmes incluant des fonctions automatisées dans certains contextes ou certaines circonstances.

3. Les aspects juridiques de l'interdiction/limitation de l'emploi.

Il convient de rappeler que pèsent sur les Etats, pour les systèmes autonomes, les obligations concernant les armes nouvelles, en particulier l'article 36 Protocole I de 1977¹³.

¹² *Idem.*

¹³ Cet article se lit comme suit: «Dans l'étude, la mise au point, l'acquisition ou l'adoption d'une nouvelle arme, de nouveaux moyens ou d'une nouvelle méthode de guerre, une Haute Partie contractante à l'obligation de déterminer si l'emploi en serait interdit, dans certaines

Mais il est clair que cette obligation est strictement procédurale. L'Etat a l'obligation d'évaluer les effets de l'arme, certes, mais cela ne peut aboutir à une quelconque prohibition, quand bien même il serait établi qu'elle est problématique.

Il ne fait guère de doutes, et cette question ne sera pas discutée, que l'arme autonome devra permettre, dans son emploi, le respect du droit international humanitaire. Même l'arme nucléaire doit, selon la Cour, y est soumise. Mais il est difficile, au-delà de cette simple constatation, d'aller au-delà. Tout d'abord, il est encore strictement impossible, au stade actuel, de conclure à une incapacité des machines à le respecter. Des doutes peuvent être émis, des interrogations soulevées, mais, quoi qu'il arrive, elles ne peuvent être tranchées par principe. On retrouve la difficulté soulevée par l'inconnu. Celui-ci barre la pensée.

Par ailleurs, il faut rappeler l'avis de la Cour internationale de justice du 8 juillet 1996. Les principes de distinction, de prohibition des maux superflus ne désignent pas des catégories d'armes interdites. Ils ne sont pas *self-executing*. Donc une arme sera interdite pour un Etat parce qu'il y aura renoncé, soit par une déclaration unilatérale, soit par la conclusion d'un traité. A ce titre, un processus d'interdiction n'est pas un absolu, mais un continuum. La conclusion d'un traité visant à interdire ou à encadrer l'autonomie est envisagée, sachant que l'on fait alors référence à des exemples plus ou moins récents: prohibition des armes à sous-munitions, mines anti-personnel. Mais on a honte de rappeler que ces traités ne contiennent que des propositions de renonciation aux Etats qui peuvent ou non lever l'option. Certes la conclusion d'un traité n'est pas indifférente pour les Etats (on l'a vu avec le traité interdiction des armes nucléaires). Mais la conclusion d'un traité actuellement ne résoudra rien dès lors que les acteurs majeurs (Russie, Etats-Unis, Chine) n'y rentreraient qu'à des conditions draconiennes.

A ce propos, on voit se développer des politiques juridiques extérieures très ambiguës, en particulier celle de la Chine, dans une position d'avril 2018. La définition d'un arme autonome se réalise par le fait qu'elle remplit tous les critères suivants, sans

circonstances ou en toutes circonstances, par les dispositions du présent Protocole ou par toute autre règle du droit international applicable à cette Haute Partie contractante»

exception aucune: lethality; autonomy, “which means absence of human intervention and control *during the entire process* of executing a task”; “impossibility for termination” such that “once started there is *no way* to terminate the device”; “indiscriminate effect,” in that it will “execute the task of killing and maiming *regardless of conditions, scenarios and targets*”; and “evolution,” “through interaction with the environment the device can learn autonomously, expand its functions and capabilities *in a way exceeding human expectations*” (emphasis added throughout).

On conçoit alors que la qualification de l’arme autonome est à ce point exigeante que son interdiction ne porterait que sur des systèmes dont la faisabilité est discutée. La portée de l’engagement serait alors extrêmement faible. Cette position n’est pas forcément isolée d’ailleurs. Celle de la France, bien que moins maximaliste, consiste également à cantonner la définition du SALA à des systèmes futurs et pleinement autonomes. Lors de la réunion du Groupe d’experts gouvernementaux (27-31 août 2018) sur la Convention sur Certaines Armes Classiques (CCAS), et à propos de la caractérisation, la France a développé les critères suivants: «les SALA sont des systèmes qui n’existent pas encore»; ils ne recouvrent pas des «systèmes automatisés ou téléopérés existant actuellement (drones, torpilles, systèmes de défense automatisés»»; «L’autonomie d’un SALA doit en effet être entendue comme complète, c’est-à-dire sans aucune forme de supervision humaine à partir de l’activation du système et une absence de subordination à une chaîne de commandement»; «Un tel système aurait la capacité de remplir des tâches dans un environnement complexe mais serait également capable de s’affranchir du cadre de règles qu’un homme lui aura confié, c’est-à-dire de configurer de nouvelles règles sans validation humaine»; «Un système qui serait capable de s’assigner lui-même des objectifs ou de modifier sans validation humaine, en cours de mission, le cadre qui lui avait été fixé pour cette mission, pourrait ainsi être considéré comme un SALA». Et la France note que de tels systèmes qui échapperaient à la supervision humaine et au contrôle du commandement «ne présentent pas d’utilité militaire». On retrouve donc ce biais définitionnel qui consiste à cantonner l’autonomie à des systèmes dont les Etats ne veulent pas. Se posera alors la question de la pertinence même d’une interdiction.

4. La dynamique de l'interdiction: réciprocité et obligations interdépendantes.

Dans le cadre de la réglementation des armements, le comportement des différents acteurs est fondamentalement mimétique. Traduit en langage juridique, ce mimétisme s'exprime par la voie des obligations interdépendantes. Cette catégorie renvoie à la nécessité d'une exécution réalisée par l'ensemble des parties tenues au respect de l'obligation, de «telle sorte qu'une violation fondamentale [de l'une] des obligations conventionnelles justifie une non-exécution correspondante par les autres parties et pas seulement une non-exécution dans les relations de ces parties avec les parties en défaut»¹⁴. Et l'interdépendance joue tant dans la formation des normes internationales, en particulier des traités, que dans leur exécution. Il s'ensuit que des engagements fermes sur ce point, et traduisant de véritables prohibitions, dépendront d'une volonté commune qui pourra être mise en échec par la mauvaise volonté d'un seul Etat. Bref, soit l'humanité renonce dans son ensemble à des armements qu'elle s'interdit, soit un Etat ou plusieurs Etats franchissent le cap, ce qui ouvre la voie à leur généralisation. Cette interdépendance correspond d'ailleurs à une triple dynamique, dont certaines sont très classiques, une autre sans doute plus originale. La première est évidente et elle est d'ordre militaire. Aucun Etat n'acceptera de se placer dans une situation d'infériorité par rapport à un adversaire potentiel ou simplement un rival. C'est une vieille leçon des choses militaires: dans les affaires de guerre, on s'aligne toujours sur le standard le plus désastreux. La deuxième est aussi classique et, comme la première, d'importance fondamentale. Elle est économique et liée aux industries de défense. Si des Etats se dotent d'industries de défense dans le domaine de l'autonomisation, de l'intelligence artificielle, les autres Etats sont forcément incités à faire de même, pour des raisons qui se cumulent mais qu'il serait inutile de développer ici. Disons seulement que se joue une part de la compétitivité des acteurs des industries de défense. Ne pas développer aboutit à perdre des avantages compétitifs.

C'est finalement la troisième dynamique qui s'avère la plus

¹⁴ V. *ACDI*, 1958, vol. II, p. 46.

originale. L'interdépendance est non seulement militaire, économique mais aussi éthique. Elle renvoie à la délimitation des adversaires auxquelles des armées pourraient être confrontées. Comme le relève G. de Boisboissel: «Dans les réflexions doctrinales américaines, il serait éventuellement plus éthique d'utiliser des systèmes robotiques autonomes pour protéger ses propres combattants. Si dans certaines situations, les soldats sont confrontés à une grave menace, selon les Américains, il faut envisager le cas où utiliser ses machines pour préserver ses propres soldats serait plus éthique»¹⁵. Th. Berthier le souligne également: «Ce qui ne serait pas éthique désormais, ce serait de ne pas utiliser ces systèmes autonomes et d'envoyer des soldats américains sur un champ de bataille devant des unités autonomisées des autres armées»¹⁶.

Ce raisonnement vise, on le voit, à la protection par les Etats-Unis de ses propres soldats contre des armes autonomes. Il traduit que l'éthique renvoie à une dimension fortement relationnelle; le comportement des autres a un impact sur l'évaluation éthique de son propre comportement. La doctrine du zéro mort, la volonté d'engager ses hommes dans des conditions moralement et opérationnellement acceptables amène à la décision d'au moins envisager de se doter de systèmes autonomes, et donc, dans l'immédiat, de ne pas se lier juridiquement. Mais on peut formuler de manière un peu plus précise les choses et qualifier le mécanisme: le droit des soldats des Etats-Unis de ne pas se faire tuer par des robots impliquera que les Etats-Unis s'en dotent. Comme il n'est pas éthique qu'un humain (états-unien) soit tué par un robot, il faut mettre des robots face aux robots.

Dans cette optique, cela signifie que la décision d'un Etat, quel qu'il soit, s'impose aux autres ou en tout cas sera susceptible de questionner leurs choix initiaux. Il fait un choix non pour lui-même, mais pour au moins ses rivaux avérés ou potentiels. Certes on peut lire aussi l'évolution de la position

¹⁵ Cité in M. BORDRON, *Armes létales autonomes: la doctrine des Etats*, dans *Actualités France Culture*, 22 novembre 2018 [<https://www.franceculture.fr/geopolitique/armes-letales-autonomes-les-doctrines-des-etats>].

¹⁶ *Ibidem*.

américaine comme un moyen de dissuader la Chine de franchir le pas. Mais peu importe finalement: le mécanisme est clair. Un Etat prendra la décision et la prendra pour tout le monde. Cela signifie qu'un Etat qui ne partage pas nécessairement nos valeurs, notre rapport à la liberté, nos conceptions éthiques, pourra imposer un nouveau standard militaire, et donc éthique/moral, à de nombreux acteurs. C'est une situation à bien des égards inédite dans l'histoire des relations militaires.

LA DECISION DU CONTRACTANT AU PRISME DE L'INTELLIGENCE ARTIFICIELLE

Garance Cattalano

SOMMAIRE: 1. Introduction. – 2 L'automatisation de la décision de conclure le contrat. – A. Les modalités pratiques. – B. Les difficultés théoriques. – 3. L'automatisation des décisions pendant l'exécution du contrat. – A. La simplification. – B. La rigidification.

1. Introduction.

Alan Turing, père de l'informatique moderne, nous posait cette question: «les machines peuvent-elles penser ?». Reprenons-la et étendons la réflexion d'un point de vue juridique: peuvent-elle contracter, ou plus largement, décider à la place d'un contractant ? Cette «pensée des machines», si elle existe, a-t-elle sa place dans notre droit des contrats moderne ?

À dire vrai, nous ne parlons pas complètement du futur: l'intelligence artificielle intervient déjà dans le processus contractuel¹. C'est le cas aujourd'hui avec l'exemple de *trading* haute fréquence, par lequel des investisseurs acquièrent et revendent en quelques secondes des actions analysées comme susceptibles de produire une marge bénéficiaire élevée. Ou encore des serrures connectées qui, au moyen d'un smart contract, bloquent automatiquement la porte d'un appartement de location si le loyer n'a pas été payé. Et à en croire les promesses des industriels, cette intervention de

¹ Pour plus de développements sur ce point, v. par ex.: M. MEKKI, *Intelligence artificielle et contrats*, dans A. BENSAMOUN, G. LOISEAU (sous la direction de), *Droit de l'intelligence artificielle*, 2019, Paris, p. 131.

l'intelligence artificielle dans le domaine contractuel sera plus prégnante encore demain, lorsque des réfrigérateurs connectés commanderont seuls les courses de la semaine en analysant les données de consommation de leur propriétaire². Il existe ainsi assurément une forme d'automatisation progressive du droit des contrats avec le développement des technologies informatiques. Or, cette automatisation pose au moins deux questions si on la voit à travers le prisme de la décision qui est le nôtre dans le cadre de ces lignes.

Pour commencer, automatisation du contrat signifie-t-il pour autant intelligence artificielle du contrat ? Car à mieux y regarder automatisme et intelligence semblent des processus radicalement opposés. D'ailleurs les automates existent depuis longtemps pour exécuter des obligations contractuelles: pensons aux distributeurs automatiques de billets ou de café dont personne n'a jamais prétendu qu'ils puissent être intelligents. Mais la frontière entre automatisation et intelligence artificielle est d'autant plus délicate à tracer que le terme «intelligence artificielle» a lui-même des contours incertains. Un rapport de la CNIL sur la question souligne d'ailleurs combien toute tentative de définition est susceptible de contestation en fonction des connaissances techniques à l'époque où la définition est proposée³. On dispose toutefois de quelques ébauches: ainsi, une norme ISO définit l'intelligence artificielle comme «*la capacité d'une unité fonctionnelle à exécuter des fonctions généralement associées à l'intelligence humaine, telle que le raisonnement et l'apprentissage*»⁴. L'intelligence, même artificielle, suppose donc la capacité à opérer un choix de façon relativement autonome. Or, pour certains auteurs, c'est précisément cette

² Sur les objets connectés v. notam.: M. BEHAR-TOUCHAIS (sous la direction de), *Les objets connectés, actes des journées du 17, 18 et 19 octobre 2017*, 2018, Paris. F. CHÉRIGNY, A. ZOLLINGER (sous la direction de), *Les objets connectés*, 2018, Poitiers.

³ CNIL, Comment permettre à l'homme de garder la main, rapp. déc. 2017: https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_garder_la_main_web.pdf, p. 14. Le rapport propose également de distinguer les IA fortes (multitâches), des IA faibles (spécialisées dans une tâche particulière) mais cette distinction tend à devenir de plus en plus poreuse.

⁴ La norme ISO 2382-28. Révisée par la norme ISO/IEC 2382:2015.

autonomie qui signe le trait caractéristique de l'intelligence artificielle et permet de la distinguer d'autres types de programmes informatiques: elle désigne ainsi « *un système informatique capable d'effectuer un choix autonome, distinct de celui de la personne qui l'a conçue ou qui en a l'usage* »⁵. L'intelligence artificielle suppose dès lors que le système opère une analyse des données pour opérer le choix le plus conforme à la volonté du programmeur, mais un choix que ce dernier n'a pu faire dès la programmation du logiciel intelligent. Une décision donc, que l'homme ne prend pas et ne maîtrise pas. C'est pourquoi, d'ailleurs, nous y reviendrons, le smart contract n'est pas un contrat intelligent au sens de l'intelligence artificielle⁶. Il s'agit en effet de l'exécution automatisée d'un droit ou d'une obligation contractuelle dont les conditions et les effets ont été prédéterminées par l'homme. Néanmoins, même le smart contract pourra un jour s'appuyer sur de l'intelligence artificielle, raison pour laquelle nous l'évoquerons largement, bien qu'il ne soit pas aujourd'hui un "contrat intelligent".

L'intelligence artificielle joue ainsi un rôle clef à toutes les étapes d'un processus décisionnel: elle permet en effet de recueillir les données nécessaires au choix, de les analyser rapidement, de les classer de les hiérarchiser voire d'opérer un choix en fonction de critères ou d'objectifs préétablis. L'intelligence artificielle avait donc toutes les raisons de devenir un outil privilégié de la prise de décision des contractants⁷, qu'elle peut rendre plus efficace, plus simple, plus rapide et donc plus économique.

Mais, et c'est là la seconde question que pose notre sujet, peut-on à proprement parler de «décisions» dans un cadre contractuel ? Parce que décider relève d'un for interne profondément unilatéral – décider c'est prendre un parti, opter pour une solution, faire un choix⁸ –, la place de la décision est à première vue limitée dans le paysage contractuel où tout

⁵ S. MERABET, *Vers un droit de l'intelligence artificielle*, thèse Université d'Aix-Marseille, 2018, n° 59 et s.

⁶ V. *infra*, IIA.

⁷ V. M. MEKKI, *Intelligence artificielle et contrats*, dans A. BENSAMOUN, G. LOISEAU (sous la direction de), *Droit de l'intelligence artificielle*, 2019, Paris, p. 131.

⁸ Dictionnaire Larousse.

se fait à deux. Le terme n'est d'ailleurs pas courant en droit des contrats: on lui préfère celui d'accord, de rencontre des volontés, ou d'échange des consentements. La lecture des pages du Code civil révèle qu'en matière contractuelle, le terme décision n'est utilisé qu'une fois, à propos d'un pouvoir unilatéral du contractant⁹. Le terme aurait ainsi, en droit des contrats, un sens limité aux prérogatives, aux pouvoirs unilatéraux dont un contractant peut user, dont il peut décider seul la mise en œuvre. Selon cette acception, la décision serait celle que l'on peut imposer à son cocontractant sans son accord préalable, ni celui du juge.

Pris dans ce sens le terme décision désigne d'ailleurs tout autant des pouvoirs de sanctions unilatérales (résolution du contrat, réduction du prix...) que des pouvoirs liés à l'exécution du contrat (modifier unilatéralement le prix initialement prévu), ou même des pouvoirs de conclusion du contrat (pouvoir de décider la reconduction ou le renouvellement d'un contrat dans des conditions déterminées à l'avance par les parties, voire matérialiser une acceptation face à une offre déjà émise). Or ces «décisions» unilatérales du contractant peuvent faire l'objet d'une automatisation. Reprenons l'exemple des objets connectés en songeant à ces boutons qui permettent de recommander automatiquement sur Amazon un produit qu'on vient de terminer. La «décision» de conclure le contrat se prend seul en appuyant sur le bouton de l'objet connecté, et même en le laissant, demain, commander seul s'il est «intelligent».

L'automatisation du contrat ou le recours à des logiciels informatiques d'aide à la décision en matière contractuelle augmente avec la contractualisation de masse, et le développement d'outils de collecte et d'analyse des données. C'est ainsi qu'aujourd'hui, les décisions en cours d'exécution du contrat – sanctionner une inexécution, fixer un prix... – (II), voire la décision de contracter elle-même (I), peuvent être prises par des programmes informatiques. Cela conduit évidemment à repenser la place de la décision et plus généralement de la volonté humaine en droit des contrats où la subjectivité est pourtant reine.

⁹ Art. 1223 C. civ.: «décision de réduire le prix»

2. L'automatisation de la décision de conclure le contrat.

L'automatisation de la décision de conclure le contrat conduit à s'interroger d'abord sur ses modalités pratiques (A), puis sur les difficultés théoriques qu'elle fait naître (B).

A. Les modalités pratiques.

L'intervention de l'intelligence artificielle dans la phase de conclusion du contrat peut prendre des formes diverses.

De nos jours, elle se manifeste dès la phase précontractuelle, au moyen notamment des outils d'aide à la décision. Mais on peut également penser aux processus de publicités ciblées, dont l'utilisation se développe avec l'accroissement de l'exploitation de nos données personnelles et plus particulièrement des données de navigation sur Internet. C'est ainsi que lors de l'acquisition de billets d'avion pour l'Italie, un volet publicitaire proposera par exemple d'acquérir un guide de conversation franco-italien. L'intelligence artificielle peut encore être utilisée pour vérifier que les documents contractuels respectent les prescriptions légales. C'est ainsi que des logiciels de reconnaissance optique des caractères peuvent être utilisés pour vérifier par exemple qu'une caution a rempli correctement la mention manuscrite qu'elle doit parfois apposer, ou que d'autres programmes se proposent de vérifier demain la conformité des clauses contractuelles au droit en vigueur grâce notamment à l'open data des décisions de justice¹⁰. D'autres logiciels préparent d'ailleurs la rédaction de contrats simples en lieu et place des rédacteurs d'actes. L'intelligence artificielle vient ici assister, ou conforter la décision de contracter prise par les volontés humaines.

Mais l'intelligence artificielle peut se faire plus présente encore: loin de simplement aider l'homme dans la prise de décision, elle peut totalement le remplacer pour la conclusion du contrat. L'intelligence artificielle sert ainsi à déclencher la

¹⁰ Comme par exemple la société LawGeex. Sur le développement de ces pratiques, v. notam. M. MEKKI, *Intelligence artificielle et contrats*, dans A. BENSAMOUN, G. LOISEAU (sous la direction de), *Droit de l'intelligence artificielle*, 2019, Paris, p. 131.

conclusion automatique d'un contrat. Et ces pratiques sont parfois anciennes: sur Ebay par exemple, le parieur prévoit une enchère maximale. Le site enchérit à la place du parieur et le contrat est conclu dès lors qu'il fait l'enchère la plus haute dans un temps donné sans dépasser le plafond fixé par le parieur. On peut encore songer à des contrats automatiquement conclus en application d'un contrat-cadre qui en définit les conditions: les objets connectés qui commandent seuls des produits, ou le *trading* haute fréquence qui, en analysant les données boursières «choisit» les ordres d'achat ou de vente les plus intéressants. Ainsi, aussi bien dans la phase précontractuelle comme outil d'aide à la décision humaine, que pour la conclusion du contrat lui-même effaçant toute manifestation de volonté humaine, l'intelligence artificielle joue un rôle de plus en plus tangible. Ce rôle remet en cause l'un des piliers de la théorie du droit des contrats: la place de la volonté. C'est dire que le développement pratique de l'intelligence artificielle nous plonge dans un abîme de difficultés théoriques.

B. Les difficultés théoriques.

Les difficultés théoriques que peut poser l'irruption de l'intelligence artificielle dans la conclusion du contrat sont importantes.

D'abord, à quoi – à qui ? – le cocontractant a-t-il affaire ? Le contrat est-il conclu par l'interposition d'une chose ou d'une personne ? Intuitivement, on y voit plutôt l'interposition d'une chose: le contrat est conclu au moyen de l'objet connecté ou du logiciel, compris alors comme un simple outil, au même titre qu'on peut utiliser un téléphone ou un ordinateur pour conclure des contrats à distance. Toutefois, l'objet dispose ici certaine forme d'autonomie et surtout son utilisateur n'a parfois exprimé aucune volonté lors de la prise de décision de contracter.

Faut-il en conclure que plus que l'interposition d'une chose, le recours à des technologies fondées sur de l'intelligence artificielle réalise en vérité l'interposition d'une forme de personne ? Pas une personne humaine évidemment, mais le droit accorde déjà la personnalité juridique à des groupement de personnes ou même de

choses ¹¹. Pourquoi, dès lors, ne pourrait-il pas accorder la personnalité juridique à une chose ? ¹² Intellectuellement admissible, la personnalité juridique de l'intelligence artificielle laisse cependant incertain sur son utilité ou son opportunité. En effet, en opportunité, en cas de difficulté, c'est l'intelligence artificielle qui répondrait de ses actes, sur son propre patrimoine... mais encore faut-il qu'il soit garni !

Et si l'on voit dans la conclusion d'une opération par le biais d'une intelligence artificielle un contrat conclu par l'interposition d'une personne, l'opération pourrait être qualifiée de mandat et risquerait à ce titre de causer de redoutables difficultés. Imaginons en effet que l'intelligence artificielle soit conçue et mise au service de l'utilisateur par le cocontractant de ce dernier, comme un objet connecté Amazon par exemple permettant de commander de nouveaux produits chez Amazon. Il y aurait là un risque élevé de conflit d'intérêt pour le mandataire (l'objet connecté) qui représente le mandant (l'utilisateur) mais a vu ses paramètres déterminés par le tiers avec qui il contracte (Amazon). Et d'autres questions se posent encore: qui informer avant la conclusion du contrat pour exécuter l'obligation précontractuelle d'information de l'article 1112-1 du Code civil ? Le contractant ou l'intelligence artificielle qui «décide» de la conclusion ? Ces informations ne doivent-elles pas lui être transmises pour qu'elle puisse les intégrer à ses critères de décision ? Il est dès lors probablement tout à la fois plus simple et plus prudent de continuer de voir l'intelligence artificielle comme une simple chose sans qu'il soit nécessaire de lui donner la personnalité juridique.

Raisonnons différemment donc, et envisageons les rapports

¹¹ A. VIAL, *La qualification juridique de l'intelligence artificielle: du silicium à la personne ?*, dans *Revue Droit & Affaires*, n° 15, Décembre 2018, 4.

¹² V. en ce sens Résolution du Parlement Européen du 16 févr. 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique (2015/2103(INL)), 59. Sur ce débat v.: G. LOISEAU, *Des droits humains pour personnes non humaines*, dans *D.*, 2011, p. 2558; G. LOISEAU, *La personnalité juridique des robots: une monstruosité juridique*, dans *JCP E*, 2018, 597; G. LOISEAU, M. BOURGEOIS, *Du robot en droit à un droit des robots*, dans *JCP G*, 2014, 1231, p. 2162. Le parlement européen est ainsi favorable à la personnalité juridique spécifique pour les robots.

entre les parties concluant des contrats par le biais d'une intelligence artificielle comme un contrat-cadre conclu entre l'utilisateur et le fournisseur, relayé par des contrats d'application dont la conclusion est déclenchée automatiquement au moyen d'un programme informatique. Mais il faut alors admettre que l'expression de la volonté de l'utilisateur est réduite: elle se résume à la conclusion du contrat-cadre. Tandis que le déclenchement des contrats d'application relève de la «décision» de l'intelligence artificielle. Ce morcellement de la volonté pourra poser différents problèmes, qui supposeront une répartition différente des risques en fonction de leur cause. Si c'est l'intelligence artificielle qui ne fonctionne pas correctement ou qui a été mal codée et qu'elle prend de mauvaises décisions, les risques doivent alors être supportés par le développeur. Mais si les mauvaises décisions de l'intelligence artificielle sont le reflet des défauts d'un contrat-cadre mal pensé, alors une répartition des risques entre les parties pourrait être envisagée. Néanmoins, il faudrait peut-être, dans cette dernière hypothèse ménager une protection spéciale du consommateur¹³. Peut-être en estimant que ne doivent être maintenus que les contrats qui préservent l'intérêt du consommateur. Ou, moins radicalement, sans doute serait-il plus judicieux d'étendre le droit de rétractation de 14 jours qui s'applique déjà aux contrats conclus à distance par un consommateur, pour les contrats conclus directement par l'intelligence artificielle. Cette dernière option permettrait au consommateur (Art. L. L221-18 C. consom.), de retrouver son pouvoir de décision... puisqu'il déciderait au moins de renoncer au contrat !

L'irruption de l'intelligence artificielle dans la sphère contractuelle soulève donc de nombreuses questions lorsqu'elle joue au stade de la conclusion du contrat, et les questions sont tout aussi nombreuses lorsque l'intelligence artificielle est utilisée pendant la phase d'exécution du contrat.

¹³ V. notam. sur cette question: S. MERABET, *Vers un droit de l'intelligence artificielle*, thèse Université d'Aix-Marseille, 2018, n° 420 et s.

3. L'automatisation des décisions pendant l'exécution du contrat.

Les décisions unilatérales en cours de contrats sont nombreuses. Elles se sont beaucoup développées, au moins dans les textes, en droit français avec la récente réforme du droit des contrats¹⁴. Ce sont essentiellement des pouvoirs de sanction des inexécutions qui dispensent désormais de faire appel à un juge (résolution unilatérale sans passer devant un juge, réduction du prix face à une exécution partielle, mise en œuvre d'une clause résolutoire etc.). Mais ces pouvoirs peuvent être des pouvoirs offerts en cours d'exécution en l'absence de toute difficulté entre les parties, comme par exemple un pouvoir de fixation unilatérale du prix en cours d'exécution dans certains contrats. À cet égard on peut tout à fait imaginer le recours à une intelligence artificielle pour fixer unilatéralement le prix d'un produit en analysant régulièrement le marché des matières premières nécessaires pour le fabriquer: l'intelligence artificielle serait ainsi utilisée pour fixer les nouveaux prix catalogues. Évidemment, le développement de ces pouvoirs de faire seul ce qui se faisait habituellement à deux ou en demandant au juge, s'accompagne d'un encadrement important afin d'éviter qu'ils deviennent l'instrument d'abus¹⁵. Précisément en raison de cet encadrement, l'automatisation de ces pouvoirs est une source de simplification pour les contractants (A) autant qu'elle aboutit à une rigidification de leur situation (B).

A. La simplification.

La simplification du jeu des décision pendant l'exécution du

¹⁴ Ordonnance n° 2016-131 du 10 février 2016 portant réforme du droit des contrats, du régime général et de la preuve des obligations, complétée par la loi du 20 avril 2018 ratifiant l'ordonnance n° 2016-131 du 10 février 2016 portant réforme du droit des contrats, du régime général et de la preuve des obligations.

¹⁵ Sur cet encadrement, v. le dossier: *Les nouveaux pouvoirs unilatéraux du contractant*, dans *RDC*, 2018, p. 503 et notam.: J.-F. HAMELIN, *L'exercice des nouveaux pouvoirs unilatéraux du contractant*, dans *RDC*, 2018, p. 514; J. HEINICH, *Le contrôle des pouvoirs unilatéraux du contractant*, dans *RDC*, 2018, p. 521.

contrat, et notamment pour mettre en œuvre les sanctions de l'inexécution, peut essentiellement provenir de l'utilisation de «smart contracts»¹⁶. Loin d'être des contrats, ceux-ci sont en fait des programmes informatiques «auto-exécutants» qui se développent comme des applications sur une blockchain: dès lors qu'une certaine condition prévue par le programmeur se réalise et que l'information parvient au programme via la blockchain, la conséquence informatique pré-codée s'exécute. Les smart contracts sont ainsi des automates numériques, des duplications informatiques d'un véritable contrat juridique, qu'on appelle le contrat «fiat». Les smart contracts peuvent ainsi coder automatiquement par exemple l'exécution du contrat, comme un contrat d'assurance qui indemnise un voyageur si son avion est en retard. Mais tout autant qu'il programme l'exécution automatisée d'un contrat, le smart contract peut en programmer l'inexécution en réaction à un défaut d'exécution de la prestation attendue et automatiser ainsi l'exception d'inexécution. L'utilisation du smart contract pourrait encore déclencher automatiquement le versement d'une clause pénale, ou le jeu d'une réduction du prix. On peut encore envisager de recourir au smart contract pour bloquer automatiquement l'accès à un appartement ou à une voiture de location (grâce à une serrure connectée) si le loyer n'a pas été payé¹⁷.

Techniquement parlant les smart contracts ne sont donc pas de l'intelligence artificielle car les conditions comme les effets sont prédéterminés par l'homme. Mais ils peuvent s'appuyer sur de l'intelligence artificielle qui par exemple analyserait les données nécessaires au déclenchement du smart contract. Corrélativement, la blockchain, utilisée comme un registre de

¹⁶ Le terme smart contract aurait été utilisé pour la première fois par Nick Szabo: N. SZABO, *Formalizing and securing public networks*, dans *First Monday*, 2, n° 9, 1er septembre 1997: <http://ojphi.org/ojs/index.php/fm/article/view/548/469>.

¹⁷ Sur les difficultés juridiques que posent les smart contracts v.: G. GUERLIN, *Considérations sur les smart contracts*, dans *D. IP/IT*, 2017, p. 512; M. MEKKI, *Le contrat objet des smart contracts*, dans *D. IP/IT*, 2019, 27; Partie I, *Le smart contract objet du droit*, *D. IP/IT*, 2019, 27; M. MEKKI, *L'avenir du droit des contrats réside-t-il dans les smart contracts ?*, dans *Mélanges offerts à G. Pignarre*, 2018, Issy-les-Moulineaux, p. 585; G. CATTALANO, *Smart contracts et droit des contrats*, dans *AJ contrat*, 2019, p. 321.

données est un outil précieux au service de l'intelligence artificielle qui peut s'appuyer sur elle pour assoir ses analyses sur des données pensées pour être fiables et sécurisées.

À cet égard, l'automatisation des décisions en cours de contrat via le smart contract présente un certain nombre d'avantages. Celui de la rapidité et de l'économie d'abord, puisque le smart contract dispense les contractants du temps et des forces nécessaires au traitement des dossiers et à l'exécution des contrats. Le smart contract permet ensuite de respecter simplement, et sans risque, les exigences formelles que la loi impose souvent pour la mise en œuvre des décisions unilatérales des contractants: mise en demeure automatique ou respect d'un délai avant le déclenchement d'une sanction, sont autant d'éléments facilement codifiables dans un smart contract.

Néanmoins la simplification apportée est loin d'être totale: les choses ne sont pas toujours si simples. Comment coder par exemple le respect d'une condition de forme si celle-ci dépend d'un standard: mettre en demeure le contractant dans un «délai raisonnable» par exemple ?¹⁸ Certes, ces affirmations doivent être nuancées car les progrès techniques vont à une vitesse considérable et l'intelligence artificielle pourra sans doute un jour s'adapter à la mise en œuvre de standards quantifiables comme le délai raisonnable en précisant dans le programme la durée du contrat, le secteur d'activité, les usages en la matière, l'importance du contrat etc. Mais, de tels programmes n'existent pas encore et surtout, si la codification des conditions de forme exigées par la loi pour le déclenchement d'un pouvoir semble envisageable, celle des conditions de fond l'est nettement moins.

Comment vérifier que le créancier a «de bonne foi» fixé unilatéralement le prix ou résolu le contrat si cette décision a été prise par le biais d'une intelligence artificielle ? Peut-il même être de bonne ou mauvaise foi dans cette hypothèse ? Car la bonne foi suppose d'écouter les explications et les difficultés du cocontractant qui ne s'est pas exécuté. Peut-être faudrait-il alors, réintroduire une décision humaine là où le smart contract tendait plutôt à l'éliminer: offrir la possibilité de contacter un préposé humain du contractant avant la mise

¹⁸ P. DE FILIPPI, A. WRIGHT, *Blockchain et droit – le règne du code*, 2019, Paris, p. 89.

en œuvre du smart contrat pour faire preuve ainsi de l'absence de mauvaise foi dans la mise en œuvre automatique d'un pouvoir de sanction ? C'est déjà toucher du doigt que cette automatisation des décisions contractuelle n'est pas toujours source de simplification¹⁹. Pire, elle peut même être source d'une difficulté supplémentaire car automatisation signifie également pétrification des rapports contractuels.

B. La rigidification.

Le propre de la volonté humaine est qu'elle est mouvante, adaptable, perméable aux influences et à la discussion, ce qu'un programme informatique est plus difficilement. Or, les décisions du contractant pendant l'exécution du contrat ont aussi besoin de cette souplesse que l'automatisation perd.

Il est ainsi très difficile de revenir sur une sanction automatiquement exécutée par un smart contrat en raison du caractère immuable des inscriptions sur la blockchain. Prévoir à l'avance le jeu d'un smart contrat codant la résolution automatique du contrat en cas d'inexécution c'est aussi se fermer au choix d'une autre sanction de l'inexécution comme par exemple la réduction du prix. Mais c'est encore se priver de la dimension négative de toute décision: décider de sanctionner c'est aussi pouvoir décider de ne pas le faire. Or, ici la décision de sanction est figée, prise à l'avance, avant même de connaître la situation qui la justifie. Certains craignent d'ailleurs que cette automatisation des décisions de sanction en matière contractuelle ne révèle un changement de paradigme contractuel, perméable à une idéologie ultra libérale «où la froideur algorithmique l'emporte sur la sagesse et la raison humaine»²⁰.

C'est ainsi que de façon très paradoxale, l'automatisation prive les décisions unilatérales des contractants d'une partie de leur utilité. En effet, ces pouvoirs unilatéraux en matière

¹⁹ J.-C. RODA, *Smart contracts, dumb contracts ?*, dans *Dalloz IP/IT*, 2018, p. 397.

²⁰ M. MEKKI, *Intelligence artificielle et contrats*, dans A. BENSAMOUN, G. LOISEAU (sous la direction de), *Droit de l'intelligence artificielle*, 2019, Paris, p. 131; G. GUERLIN, *Considérations sur les smart contracts*, dans *D. IP/IT*, 2017, p. 512.

contractuelle se sont développés essentiellement parce qu'on a pris conscience que tout ne pouvait pas être décidé dès la conclusion du contrat, qu'il fallait pouvoir faire des choix en s'adaptant aux circonstances pour introduire plus de souplesse dans la relation contractuelle. Or, l'automatisation oblige les parties à prévoir dès la conclusion du contrat des modalités de déclenchement de leurs pouvoirs unilatéraux, à figer leurs rapports au jour même où ceux-ci naissent, et les conduit donc à se priver d'une partie de leur pouvoir de décision.

Pour conclure, l'intelligence artificielle est déjà dans le contrat et refoule la volonté humaine dans un domaine où elle était pourtant reine. Cette automatisation est utile et efficace pour les contrats de masse aux prestations simples, mais semble encore trop balbutiante pour suppléer l'homme lorsque les prestations sont plus complexes ou que le sentiment retrouve sa place en droit des contrats. L'intelligence artificielle est donc assurément l'avenir de la décision en droit des contrats, mais elle ne sera pas son unique avenir.

LA LEGITIMITE DES DECISIONS CONTRACTUELLES EMANANT D'UNE INTELLIGENCE ARTIFICIELLE

Frédéric Dournaux

Sommaire: 1. Introduction. – 2 La prise de décision contractuelle par une intelligence artificielle. – A. L'intelligence artificielle, aide à la décision contractuelle. – B. L'intelligence artificielle, acteur de la décision contractuelle. – 3. De quelques questions posées par la prise de décision contractuelle par une intelligence artificielle. – A. Les insuffisances de la technique. – B. Les risques d'externalisation. – C. Les réserves d'ordre axiologique.

1. Introduction.

A l'inverse de la pensée chinoise qui privilégie le non-agir, la transformation insensible, notre civilisation occidentale valorise la décision depuis toujours, l'action héroïque et résolue, le choix décisif qui permet de forcer la fatalité. Œdipe qui se crève les yeux, César qui franchit le Rubicon, Napoléon qui force la victoire à Arcole: aucun d'eux n'hésite, tous savent décider. Au contraire, l'indécision est une tare, la marque d'un esprit faible. C'est Hamlet qui ne se décide pas à venger son père, ni à se suicider et qui ne sait même pas décider entre «*être ou ne pas être*» ! C'est encore l'âne de Buridan qui, ne sachant se décider entre boire ou manger, meurt de soif et de faim. L'étymologie du mot «*décision*» est éclairante: c'est un dérivé du latin *caedere*, qui signifie «*couper* ». Décider, c'est trancher, couper le fil des événements en imposant son choix.

En droit privé français, le terme de décision n'a pas de sens spécifique. On parle au sens courant de décisions de justice, notamment dans le Code civil, ou de décisions d'une collectivité ou d'un organe comme, par exemple, les décisions des associés d'une société ou d'un conseil de famille. Si on resserre le faisceau sur le droit des contrats, il apparaît que la

décision est partout et nulle part. Nulle part, en apparence, car le terme de décision n'est pas expressément utilisé en droit des contrats. Partout, en réalité, car le contrat n'est que la somme des décisions des parties¹. Or, on constate que la prise de décision contractuelle est de plus en plus appelée à être effectuée par une «intelligence artificielle». Il faut entendre par là toutes les techniques ayant la capacité d'«exercer des fonctions généralement associées à l'intelligence humaine, telles que le raisonnement ou l'apprentissage»². Concrètement, il s'agit de technologies très diverses à base d'algorithmes. On parle essentiellement d'auto-apprentissage (*machine learning*), d'apprentissage profond (*deep learning*) ou encore de *cognitive computing* pour désigner la modélisation de la cognition humaine (c'est par exemple le cas du *Human brain project* européen). La *Blockchain* ne relève pas en soi de l'intelligence artificielle mais elle peut y apporter son concours pour permettre le stockage et la transmission sécurisée d'informations authentifiées. Même chose pour le *cloud* qui n'est pas seulement une solution de stockage mais permet aussi de diffuser auprès de très nombreux utilisateurs des technologies d'intelligence artificielle à travers des solutions de *cloud computing*. A vrai dire, on ne peut guère circonscrire précisément pour l'heure l'intelligence artificielle. Tout au plus peut-on en distinguer deux intensités différentes: l'intelligence artificielle dite faible simule ponctuellement un processus intelligent spécifique alors que l'intelligence artificielle dite forte vise à produire une intelligence autonome générale.

Ces techniques se développent rapidement dans la pratique du droit des contrats. D'où cette double interrogation: de quelle manière l'intelligence artificielle va-t-elle modifier la prise de décision en matière de contrats et quel regard peut-on jeter sur cette révolution technique et juridique ? Pour tenter d'y répondre, nous pouvons procéder en deux temps. Dans une première partie, nous présenterons les différentes manières dont l'intelligence artificielle affecte la prise de décision contractuelle (2). Dans une deuxième partie, nous

¹ Cf. la contribution de G. Cattalano quant aux décisions contractuelles.

² Norme de certification ISO 2382:2015.

envisagerons quelques unes des questions que suscite la prise de décision contractuelle par une intelligence artificielle (3).

2. La prise de décision contractuelle par une intelligence artificielle.

De manière indirecte, l'intelligence artificielle peut être une aide à la décision contractuelle (A). De manière directe, l'intelligence artificielle peut être acteur de la décision contractuelle (B).

A. L'intelligence artificielle, aide à la décision contractuelle.

Au stade de la formation du contrat, l'intelligence artificielle est utile pour apprécier l'opportunité de contracter. Plusieurs aspects sont concernés.

L'intelligence artificielle peut d'abord inciter à contracter en identifiant un éventuel besoin à satisfaire. Certaines banques d'affaires l'utilisent pour proposer des rapprochements d'entreprises ou pour déterminer des cibles potentielles à acquérir. Suivant la même idée, la compagnie d'assurance contextuelle *Neosurance* identifie le contexte où se trouve une personne (par exemple dans un hall d'aéroport) et lui propose un produit d'assurance ponctuel (par exemple l'assurance perte bagage pour le trajet).

L'intelligence artificielle permet aussi de mieux décider grâce à l'analyse des risques contractuels. L'analyse des risques, notamment les *due diligence*, est systématique avant tout contrat d'importance. L'intelligence artificielle permet d'analyser en très peu de temps des milliers de documents et de repérer les éléments à signaler à un vérificateur humain. La société *Drooms*, par exemple, est le leader européen des *data-rooms* virtuelles.

L'analyse des risques peut aussi porter sur la personne du cocontractant. Plusieurs solutions d'intelligence artificielle sont déjà utilisées par les établissements financiers ou d'assurance pour le *scoring* de leurs clients. La note attribuée détermine la décision de leur prêter ou de les assurer et les conditions contractuelles à leur appliquer.

Outre l'analyse des risques, l'intelligence artificielle permet aussi une meilleure analyse des coûts de transaction. Les

économistes ont développé l'idée que les personnes arbitrent entre plusieurs alternatives contractuelles essentiellement en considération des coûts induits par le contrat. Aussi, des logiciels de *pricing* visent à faciliter cette décision par l'analyse de données synthétisées par intelligence artificielle. C'est par exemple le cas de *Geophy* qui prétend déterminer la valeur objective exacte d'un bien immobilier commercial ou résidentiel.

L'intelligence artificielle, par ailleurs, peut aider à l'analyse et à la rédaction des clauses du contrat. L'analyse des clauses – *contract review* – permet de repérer les risques. *eBrevia*, *Luminance* aux Etats-Unis, *Hyperlex* et *SoftLaw* en France proposent de tels produits. C'est le cas aussi de *LawGeex* connue pour avoir confronté son intelligence artificielle à 20 juristes expérimentés. Le défi consistait à identifier 30 types de clauses dans cinq accords de confidentialité (*non disclosure agreement*). L'intelligence artificielle a mis 26 secondes contre 1 h 30 en moyenne pour les humains, et a obtenu un taux de réussite de 94% contre 85% pour les humains³. L'intérêt est de faire gagner du temps pour reporter l'expertise humaine sur les clauses plus délicates.

En revanche, la rédaction automatisée de contrats est encore embryonnaire. A vrai dire, il s'agit pour l'heure d'une personnalisation de contrats types, parfois à l'aide de questionnaires menés par des assistants robots virtuels, des « *chatbots* ».

Une autre voie est essayée, celle de l'écriture numérique des contrats. Le contrat est directement écrit en code afin de l'intégrer nativement dans le système informatique pour sa gestion et son exécution et pour alimenter la base de données contractuelles qui servira pour l'apprentissage de la production de contrats par l'intelligence artificielle.

Enfin, l'aide à la décision contractuelle peut concerner l'exécution du contrat. L'intelligence artificielle n'y est pas encore très présente mais est appelée à s'y développer: les mêmes produits qui, en amont, analysent déjà l'intérêt à contracter pourront assurément, en aval, analyser la valeur des options pouvant être décidées en cours de contrat. On songe

³ B. DONDERO - B. LAMON, *Juristes humains contre IA: l'analyse de contrats, A` propos de l'étude LawGeex*, dans *JCP G*, 2018, n° 1201.

en particulier à la possibilité de ne pas exécuter le contrat s'il s'avère que cette violation est «efficace» au sens où elle procure un profit supérieur à l'exécution. L'intelligence artificielle pourra aider à décider de cette *efficient breach of contract* en comparant les coûts respectifs de l'exécution et de l'inexécution, notamment à l'aide des données de justice prédictive ou au regard du coût de remplacement par un autre contrat.

Franchissons maintenant une étape supplémentaire: d'assistant à la décision contractuelle, l'intelligence artificielle peut devenir acteur de la décision contractuelle.

B. L'intelligence artificielle, acteur de la décision contractuelle.

Il y a là deux aspects à évoquer: l'intelligence artificielle peut être décideur contractuel pour autrui; mais on peut aussi envisager que, dans le futur, elle devienne un décideur contractuel autonome.

L'intelligence artificielle, décideur contractuel pour autrui, cette situation existe déjà. Des systèmes informatiques peuvent être programmés pour prendre seuls des décisions contractuelles, sans intervention humaine. Une machine peut ainsi décider de conclure un contrat dont les conditions prédéterminées sont réunies. Tel est le cas du trading algorithmique qui représenterait 80 % des transactions financières dans le monde. Le système informatique décide seul, en fonction de centaines de paramètres, d'acheter ou de vendre, tel produit financier, à tel prix, pour tel volume, à tel moment, sur tel marché, à telles conditions...

L'intelligence artificielle peut aussi décider d'exécuter un contrat déjà conclu. L'exécution mécanique est ancienne: c'est déjà le cas d'un simple distributeur de boissons. Nul besoin d'intelligence. Il en va de même des mal nommés *smart contracts* qui ne sont ni intelligents, ni des contrats, juste des logiciels d'automatisation de l'exécution d'un contrat. La *blockchain* en multiplie les usages mais ne les rend pas plus intelligents. Ils jouent leur partition comme un piano mécanique: transferts de fonds, communication de documents, gestion du début ou de la fin d'un contrat comme, par exemple, le verrouillage à distance d'une serrure. A l'avenir, l'intelligence artificielle devrait leur donner de nombreuses

applications. On le voit déjà avec l'assureur *Lemonade* qui traite les demandes des assurés par une intelligence artificielle qui, en moins de 10 secondes, décide de l'indemnisation du client.

Le dernier stade du développement de l'intelligence artificielle en droit des contrats sera celui de l'intelligence artificielle, décideur contractuel autonome. Dans ce cas, l'intelligence artificielle n'agit plus pour autrui mais pour elle-même. Elle est autonome, au sens littéral de *auto-nomos* «*qui se régit par ses propres lois*». Evidemment, pour être partie à un contrat, il faut une capacité de contracter, un patrimoine, autrement dit une personnalité juridique. Qu'à cela ne tienne, plusieurs initiatives visent à reconnaître une troisième personne, ni physique, ni morale. Dans une résolution adoptée en 2017, le Parlement européen s'est exprimé en faveur de la création d'une personnalité juridique spécifique à tout robot qui prend des décisions autonomes ou qui interagit de manière indépendante avec des tiers⁴.

On le voit, nous sommes au début d'une ère nouvelle: l'intelligence artificielle va bouleverser la manière de décider en droit des contrats. Cette *disruption* suscite des interrogations légitimes; nous allons maintenant, dans une deuxième partie, en évoquer quelques unes.

3. De quelques questions posées par la prise de décision contractuelle par une intelligence artificielle.

Je n'évoque ici que trois séries de réflexions qui concernent respectivement la technique (A), les risques liés à l'externalisation (B) et les valeurs sous-jacentes de cette évolution (C).

⁴ Parlement européen, Résolution du 16 févr. 2017. Adde G. LOISEAU, *La personnalité juridique des robots: une monstruosité juridique*, dans *JCP G*, 2018, 597; A. BENSAMOUN - G. LOISEAU, *L'intégration de l'intelligence artificielle dans l'ordre juridique en droit commun: questions de temps*, dans *Dalloz IP/IT*, 2017, p. 239; en faveur de cette solution V. not. A. BENSOUSSAN, *Plaidoyer pour un droit des robots: de la «personne morale» à la «personne robot»*, dans *La Lettre des juristes d'affaires*, 23 oct. 2013, n° 1134.

A. Les insuffisances de la technique.

L'intelligence artificielle demeure d'une efficacité très limitée en tant qu'aide à la décision contractuelle. Pour l'heure, il y a plus d'artificialité que d'intelligence. Le marketing et les médias donnent une image déformée des capacités réelles. Les succès aux échecs ou au jeu de go relèvent plus du phénomène de foire que d'une forme d'intelligence. L'intelligence artificielle ne donne des résultats efficaces que lorsqu'elle est parfaitement paramétrée. En dehors des cadres, elle demeure assez stupide. La société *Vicarious* a montré qu'une intelligence artificielle imbattable au jeu du casse-briques est incapable de s'adapter dès lors qu'on change la hauteur ou la taille de la raquette.

Plus précisément, les deux techniques phares de l'intelligence artificielle mettront beaucoup de temps à produire leurs effets en droit des contrats. Il s'agit d'abord du *machine learning*. L'auto-apprentissage suppose une quantité considérable de données, de « *big data*»; il faut du grain à moudre pour qu'il fonctionne; les données sont ici plus importantes que l'intelligence qui les traite. Or, en droit des contrats, les échantillons pour enrichir la base sont structurellement faibles. Les contrats complexes demeurent confidentiels; les modèles contractuels originaux d'un cabinet d'affaires sont jalousement gardés. Seuls les contrats basiques sont répandus mais ils n'ont pas de valeur ajoutée. Se pose ainsi le problème du *small data*. Ce n'est pas un hasard si l'étude précitée de *LawGeex* portait sur des accords de confidentialité: c'est qu'ils sont les plus courants et les plus standardisés !

Justement, l'autre technique qui devra faire ses preuves est celle du *natural language processing*, le traitement du langage naturel. Le droit est un langage technique dont l'accessibilité formelle ne correspond pas à l'intelligibilité substantielle. L'intelligence artificielle peut identifier des mots mais ne pas en saisir le sens. Il faut pour cela maîtriser la grammaire et la sémantique du droit. Ainsi que l'écrivait le Doyen Cornu « *la polysémie est une marque essentielle du vocabulaire juridique. [...] En droit, le nombre des signifiés est infiniment plus élevé que celui des signifiants,*

les notions juridiques beaucoup plus nombreuses que les mots pour les nommer»⁵.

Indépendamment de cette polysémie, l'intelligence artificielle butera aussi sur la compréhension des notions floues, de plus en plus nombreuses en droit des contrats. Comment pourra-t-elle comprendre la bonne foi, le légitime, le raisonnable, l'abus, l'état de dépendance, le manifestement excessif, le suffisamment grave, toutes notions à contenu variable, dont la signification est contextuelle et qui supposent non seulement « *l'esprit de géométrie* » comme disait Pascal mais aussi « *l'esprit de finesse* » lesquels, pour l'heure, sont encore l'exclusivité des humains (et encore pas de tous) ?

B. Les risques relatifs à l'externalisation.

L'intelligence artificielle exige le recours à un tiers extérieur. En plaçant l'intelligence artificielle au cœur des décisions contractuelles, on y place dans le même temps les entreprises technologiques. Plus l'intelligence artificielle sera sophistiquée, plus elle se développera en *cloud computing*.

Cela signifie d'abord un risque de dépendance vis-à-vis des *legaltech* détentrices de ces capacités informatiques, risque d'autant plus grand que l'automatisation de ces tâches aura conduit à une perte en interne de ces mêmes compétences et des moyens humains dédiés.

Il en résultera aussi un risque important pour la confidentialité des décisions contractuelles. Tout recours à un tiers entraîne un risque de perte de contrôle des données confidentielles, non seulement de celles directement contractuelles mais aussi de toutes les informations qui auront été fournies à l'intelligence artificielle pour aider à la décision⁶. A cet égard, il faut rappeler que les principaux acteurs du *cloud* sont ceux-là mêmes qui font commerce du *big data*... Ne pourront-ils pas exploiter des données certes anonymisées mais néanmoins sensibles ? Et *quid*, par ailleurs,

⁵ *Linguistique juridique*, dans D. ALLAND, S. RIALS (dir.), *Dictionnaire de la culture juridique*, 2003, Paris, p. 954.

⁶ A. BENSAMOUN - G. LOISEAU, *L'intégration de l'intelligence artificielle dans certains droits spéciaux*, dans *Dalloz IP/IT*, 2017, p. 295.

du risque de vol de données confidentielles par des hackers ou des employés ? Le risque n'est pas minime dès lors que le recours au *cloud computing* entraîne la circulation des données entre plusieurs intervenants extérieurs au lieu du huis-clos des cabinets d'avocats, tenus au secret professionnel et à des règles déontologiques. Le *cloud act* du 23 mars 2018 renforce ces craintes puisqu'il permet aux autorités américaines d'exiger des prestataires de *cloud*, dans le cadre d'une procédure pénale, qu'ils divulguent les données demandées et ce, en quelque endroit du monde où se trouvent les serveurs qui les stockent⁷.

Enfin, une dernière série d'interrogations concerne les valeurs inhérentes à la prise de décision contractuelle par une intelligence artificielle.

C. Les réserves d'ordre axiologique.

Une première réserve tient à l'idéologie sous-jacente. L'intelligence artificielle n'est pas neutre, elle n'est pas seulement un outil. Elle procède d'une conception libérale qui promeut le dépassement de l'humain trop lent, trop coûteux, trop faillible, trop peu productif, pas toujours rationnel, ni même toujours docile... En droit des contrats, l'intelligence artificielle est le cheval de Troie de l'analyse économique du droit. Elle y véhicule des valeurs de rentabilité, de réduction des coûts, d'optimisation des processus, de désintermédiation.

Mais l'intrusion d'une logique économique n'est pas forcément chose facile. Le droit a sa logique propre et ses mécanismes peuvent prévaloir sur les intérêts boutiquiers. Reprenons l'exemple précité de l'*efficient breach of contract*. L'intelligence artificielle pourra démontrer à un contractant qu'il est avantageux de ne pas exécuter le contrat. Rationnelle sur le plan économique, cette décision ne l'est pas forcément sur le plan juridique. D'abord parce que le droit français ouvre au créancier plusieurs remèdes (art. 1217 du Code civil), dont l'exécution forcée en nature (1221). Ensuite parce que l'inexécution délibérée du contrat, violation du devoir de

⁷ V. not. F. PLÉNACOSTE - E. DAUD, *Cloud Act: des inquiétudes légitimes*, dans *Dalloz IP/IT*, 2018, p. 680.

bonne foi, constitue une faute dolosive: la responsabilité du débiteur défaillant n'est alors plus limitée au préjudice prévisible et l'éventuelle clause limitative de responsabilité sera tenue en échec. De quoi sans doute déjouer le calcul d'une intelligence artificielle...

Cet exemple illustre aussi un biais conceptuel à ne pas sous-estimer: pour l'heure, l'intelligence artificielle juridique est pensée pour la *Common law*, c'est-à-dire pour un système dans lequel la jurisprudence est essentielle et la doctrine négligeable. Dans la tradition civiliste, l'intelligibilité du droit vient encore de la doctrine et la jurisprudence n'est qu'une donnée brute qui doit être ordonnée et interprétée. Aussi la compréhension profonde d'un système juridique – dans ses finalités propres, ses spécificités nationales, ses équilibres, ses tréfonds historiques, moraux et sociologiques – sera-t-elle encore durablement réservée à l'intelligence humaine.

Ce qui nous amène à une seconde réserve axiologique, par laquelle je conclurai: le risque de voir l'intelligence artificielle refouler la nature humaine du contrat. Pour paraphraser Fouillée « *Qui dit contractuel dit humain* ». Le contrat est une institution pensée par et pour des humains, même lorsque ce sont des personnes morales qui le concluent. Le contrat est humain par les volontés qui le créent et en façonnent le contenu. Le droit français des contrats est d'ailleurs réputé très psychologique, attentif aux vices du consentement, aux motivations profondes des parties (avec la cause et désormais le but du contrat), avec l'exigence de bonne foi qui innerve de nombreuses solutions concrètes.

L'irruption d'une machine, même intelligente, risque de rejeter l'humain à la marge du contrat⁸. Pour l'heure, l'intelligence artificielle s'impose dans les tâches rébarbatives et à faible valeur ajoutée. Mais au fur et à mesure de ses progrès, son expertise pourrait évacuer l'humain de la réalité du pouvoir de décision et le réduire au rôle de simple signataire de ce qui, par hypothèse, a été mathématiquement décidé pour maximiser l'utilité qu'il retirera du contrat. Ce bannissement de l'humain, moins performant que la machine, est déjà à l'œuvre dans le *trading* algorithmique: aucun

⁸ G. CHANTEPIE, *Le droit en algorithmes ou la fin de la norme délibérée ?*, dans *Dalloz IP/IT*, 2017, p. 522.

humain ne peut prendre des décisions en microsecondes, c'est-à-dire en 1 millionième de seconde. Ce bannissement de l'humain est aussi en germe dans l'idée de personnalité juridique des robots. Les faire accéder au statut de personne permettra de leur reconnaître des droits humains suivant la même dérive anthropomorphique que celle déjà à l'œuvre pour les personnes morales. Mais surtout, en établissant une équivalence entre les choses et les humains, on parviendra à une concurrence des formes d'intelligence, puis à la marginalisation de l'humain par le constat de la supériorité de l'intelligence artificielle pour prendre des décisions contractuelles optimales.

Quelles seront alors ces décisions contractuelles en l'absence de véritable intervention humaine ? Il s'agira de contrats juridiquement impeccables, dénués d'incertitude et économiquement profitables. Mais cela aura un prix: la liberté contractuelle pourrait dépérir, cette liberté d'inventer des contrats *sui generis* ou de concevoir des clauses véritablement sur-mesure. L'intelligence artificielle est stérile, elle ne crée pas car elle apprend du passé et n'est pas douée d'inventivité. Elle se bornera donc aux types de contrats qu'elle a appris et aux clauses qui, par l'inlassable répétition du *machine learning* sur les mêmes bases, deviendront des normes contractuelles, des contrats d'adhésion personnalisés. Pour demeurer une institution vivante, le contrat a besoin de l'humain avec son intuition, sa sensibilité, sa perception du monde et des autres, sa capacité à oser, à compatir, à prendre des risques, à affronter l'imprévu, voire à être irrationnel ou à servir des valeurs non-marchandes...

Notre siècle est celui du prestige technicien; notre époque s'incline devant la supériorité de la science et la vérité prétendument définitive des chiffres. Mais le droit n'est pas une science, c'est un art: «*Jus est ars boni et aequi* » affirmait Celse il y a 2000 ans. Une intelligence artificielle pourra-t-elle comprendre un jour ce qui est juste et bon ? Vous l'aurez compris, à mon sens, l'intelligence artificielle est un bon serviteur mais un mauvais maître.

DRIVERLESS CARS:
I NUOVI CONFINI DELLA RESPONSABILITÀ CIVILE
AUTOMOBILISTICA E PROSPETTIVE DI RIFORMA

Stefano Pollastrelli

SOMMARIO: 1. Premessa. – 2. Dal conducente al robot. La categoria dei *Self Driving Vehicles* – 3. Il crepuscolo dell’art. 2054 c.c. Verso una nuova visione della responsabilità civile automobilistica – 4. Considerazioni conclusive.

1. Premessa.

La diffusione delle diverse applicazioni di tecnologia robotica in ambiti sempre più comuni della società civile e l’introduzione sempre più invasiva dell’automazione nel settore industriale, e nel suo interno nei diversi processi produttivi, ed in quello commerciale, hanno dischiuso problematiche di natura giuridica (ma anche di natura sociale ed etica) in quanto vi sono plurimi settori (giustizia, medicina, chirurgia, comunicazione, agricoltura, logistica, reti stradali, trasporti, ...) che sono stati fortemente incisi, rendendo vacillante parte degli impianti tradizionali normativi. Appare, pertanto, necessario muovere da una rivisitazione delle costruzioni giuridiche dei diversi settori compenetrati dall’impatto dell’intelligenza artificiale¹.

L’indagine giuridica delle intersezioni tra intelligenza artificiale e diritto ha coinvolto estensivamente il settore dei trasporti (più in generale quello della mobilità)² con l’obiettivo finale e

¹ Cfr S. CRISCI, *Intelligenza artificiale ed etica dell’algoritmo*, in *Foro amm.*, 2018, p. 1787 ss.; L. BUTTI, *Auto a guida autonoma: sviluppo tecnologico, aspetti legali ed etici, impatto ambientale*, in *Riv. giur. ambiente*, 2016, p. 435 ss.; E. PALMERINI, *Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea*, in *Resp. civ. prev.*, 2016, p. 1816 ss.

² In argomento, si vedano: L. AMMANNATI - A. CANEPA, *Politiche per un*

primario di renderlo sempre più sganciato dal fattore “uomo” (pilota, conducente, comandante di nave e di aeromobile, equipaggio) attraverso l’impiego di veicoli completamente autonomi. Nel contesto, se da un lato la navigazione aerea è, oramai, da tempo caratterizzata dall’impiego dell’intelligenza artificiale applicata agli aeromobili a pilotaggio remoto (senza la presenza di equipaggio e pilota a bordo)³, dall’altro, la scienza giuridica si sta occupando in maniera approfondita recentemente dell’impiego dell’intelligenza artificiale al settore dell’*automotive* (e cioè dei veicoli a guida automatica), che ha già ricevuto, come di seguito si vedrà, particolari attenzioni da parte del legislatore di Paesi stranieri.

Si tratta indubbiamente di un rinnovamento straordinario nel mondo del trasporto stradale, ed in particolare nel settore della circolazione stradale, e sembrerebbe anche repentino e ne è dimostrazione, in tal senso, la “Dichiarazione sulla cooperazione nel campo della guida connessa ed automatizzata” (cosiddetta “Dichiarazione di Amsterdam”), siglata nel 2016 dai Ministri dei trasporti dei Paesi membri dell’Unione europea, volta a facilitare l’introduzione e la commercializzazione di veicoli completamente automatizzati a partire dal 2019 sui mercati automobilistici⁴.

trasporto sostenibile, Napoli, 2017; G. GASPARI, *L’acquis dell’Unione europea in materia di sicurezza stradale e sistemi di trasporto intelligente*, in *Riv. dir. trasp.*, 2019, p. 1 ss.

³ In materia, cfr. A. MASUTTI, *Prospettive di regolamentazione dell’uso dei veicoli senza pilota (UAV) nello spazio comune*, in *Dir. trasp.*, 2007, p. 783; U. LA TORRE, *Gli UAV: mezzi aerei senza pilota*, in *Sicurezza, navigazione e trasporto*, a cura di R. TRANQUILLI - LEALI e E.G. ROSAFIO, Milano, 2008, p. 93 ss.; E.G. ROSAFIO, *Considerazioni sui mezzi aerei a pilotaggio remoto e sul regolamento ENAC*, in *Riv. dir. nav.*, 2014, p. 787 ss.; B. FRANCHI, *Gli aeromobili a pilotaggio remoto*, in *Resp. civ. prev.*, 2014, p. 1770 ss.; A. ZAMPONE, *Riflessioni in tema di responsabilità nell’esercizio di Remotely-Piloted Aircraft System (RPSA)*, in *Dir. trasp.*, 2015, p. 63 ss. Anche il comparto marittimo rivolge particolare attenzione allo studio delle *unmanned ships* (navi senza equipaggio - a pilotaggio remoto). Sul punto, cfr. M. CHWEDCZUK, *Analysis of the Legal Status of Unmanned Commercial Vessels in U.S. Admiralty and Maritime Law*, in *Jour. Mar. Law & Comm.*, 2016, p. 123 s.; G. BOI, “*Navi-drone*”: *primi interrogativi in tema di disciplina giuridica*, in *Riv. dir. nav.*, 2017, p. 175 ss.; U. LA TORRE, *Navi senza equipaggio e shore control operator*, in *Riv. dir. trasp.*, 2019, p. 487 ss.

⁴ L’impegno raggiunto è quello di sviluppare un quadro normativo europeo armonizzato al fine di consentire l’uso transfrontaliero delle auto connesse, nel rispetto delle medesime regole sia sotto il profilo dell’accesso ai dati, sia sotto quello della sicurezza stradale. La Dichiarazione di Amsterdam, difatti,

In questa nuova cornice, è emerso, da particolari studi di settore, che l'impiego della tecnologia fondata sull'intelligenza artificiale ridurrebbe drasticamente il numero delle vittime della strada e comporterebbe un beneficio in termini di costi assicurativi oltre che una riduzione dell'inquinamento ambientale.⁵ Ciò nondimeno, se è vero che il 90% degli incidenti stradali sono attribuibili al comportamento umano (stanchezza, sonnolenza, distrazione, stato di ebbrezza da alcool, stato di alterazione psicofisica per uso di sostanze stupefacenti o psicotrope, ...), è anche ugualmente vero che si possono profilare rischi di sinistri stradali a causa di disfunzioni nella tecnologia installata, cattivo funzionamento del *software* di bordo, guasti nella rete satellitare, errori da parte di algoritmi e manomissioni illecite di terzi ed attacchi informatici (*cybersecurity*)⁶, in sostanza, vizi e difetti della *driverless cars technology*

Dunque, lo sviluppo del settore tecnologico automobilistico dovrebbe procedere contestualmente con lo sviluppo tecnologico delle reti ed infrastrutture stradali (le cosiddette *smart road*) al fine di creare un reciproco interscambio di informazioni e dati tra infrastrutture e veicoli di nuova generazione volto alla conoscenza dei flussi e della gestione del traffico (in caso di incidenti o gravi ostruzioni), alla sicurezza stradale, ai suggerimenti di corsie⁷. L'Italia, con Decreto Ministeriale 28 febbraio 2018, ha avviato la trasformazione digitale del patrimonio infrastrutturale stradale esistente a supporto dei veicoli connessi e la sperimentazione di immettere su strade pubbliche veicoli a guida automatica dietro apposita autorizzazione ministeriale⁸.

pone importanza alla interoperabilità dei servizi e sistemi tra veicoli e tra veicoli ed infrastrutture di comunicazione.

⁵ Studio condotto da KPMG nel 2017 (*White Paper*) "*The caothic middle, the autonomous vehicle and disruption in automobile insurance*". Sul punto, cfr. P. SHARON, *Autonomous vehicle liability, insurance and regulation*, in *Dir. mer. assic. fin.*, 2017, p. 441 ss.

⁶ Questione ampiamente dibattuta in campo aeronautico. Sul punto, F. SALERNO, *Nuove frontiere della security: la cybersecurity*, in *Riv. dir. nav.*, 2018, p. 555 ss.

⁷ Cfr. S. SCAGLIARINI, *Smart roads e driverless cars nella legge di bilancio: opportunità e rischi di un'attività economica "indirizzata e coordinata a fini sociali"*, in *Quaderni cost.*, 2018, p. 497 ss.

⁸ Si tratta del d.m. 28 febbraio 2018 (Ministero delle infrastrutture e dei trasporti), concernente "Modalità attuative e strumenti operativi della sperimentazione su strada delle soluzioni di *Smart Road* e di guida connessa e automatica" (in G.U. n. 90 del 18 aprile 2018). Per *Smart Road* si intendono "le infrastrutture stradali per le quali è compiuto, secondo le specifiche funzionali

Lo scenario superiormente descritto, che come si ripete, è proiettato verso la completa automazione dei veicoli ammessi alla circolazione stradale senza conducente, pone una serie di riflessioni giuridiche in materia di responsabilità civile, assicurativa ed anche in quella più propriamente specifica della circolazione stradale⁹.

In questa fase, caratterizzata dalla veloce dinamica dell'innovazione tecnologica che condurrà presto il mercato automobilistico alla *full automation*, l'indagine da intraprendere va condotta verso la disamina della tenuta delle tradizionali categorie giuridiche incasellate all'interno del codice civile, del codice delle assicurazioni e del codice della strada, che inevitabilmente risentono delle fisiologiche fibrillazioni dell'impatto della robotica.

di cui all'art. 6, comma 1, un processo di trasformazione digitale orientato a introdurre piattaforme di osservazioni e monitoraggio del traffico, modelli di elaborazione dei dati e delle informazioni, servizi avanzati ai gestori delle infrastrutture, alla pubblica amministrazione e agli utenti della strada, nel quadro della creazione di un ecosistema tecnologico favorevole all'interoperabilità tra infrastrutture e veicoli di nuova generazione" (art. 2, comma 1). Sull'argomento, cfr. D. CERINI, *Dal Decreto Smarts Road in avanti: ridisegnare responsabilità e soluzioni assicurative*, in *Danno e resp.*, 2018, p. 405 ss.; L. BUTTI - F. RIGO - L. TRONCONI, *Decreto "smart road" quali profili di sicurezza?*, in *ambientesicurezza.it*, 2018, p. 34 ss.

⁹ Cfr. M.C. GAETA, *Automazione e responsabilità civile automobilistica*, in *Resp. civ. prev.*, 2016, p. 1718 ss.; A. DAVOLA - R. PARDOLESI, *In viaggio col robot: verso nuovi orizzonti della r.c. auto ("driverless")?*, in *Danno e resp.*, 2017, p. 616 ss.; V. DE PALMA, *La auto a guida autonoma e la responsabilità civile nella circolazione stradale*, in *Diritto.it*, 2017, p. 1 ss.; C. SEVERONI, *Prime considerazioni su un possibile inquadramento giuridico e sul regime di responsabilità nella conduzione dei veicoli a guida autonoma*, in *Riv. dir. trasp.*, 2018, p. 331 ss.; G. VOTANO, *La responsabilità da circolazione stradale nella fase di transizione dai veicoli tradizionali alle auto a guida autonoma*, in *Danno e resp.*, 2019, p. 330 ss.; A. VEDASCHI - G.M. NOBERASCO, *Gli autoveicoli a guida autonoma alla prova del diritto*, in *Dir. pubbl. comp. eur.*, 2019, p. 769 ss.; S. PELLEGGATA, *Autonomous Driving and Civil Liability: The Italian Perspective*, in *Riv. econ. trasp. amb.*, 2019, p. 135 ss.; G. BENELLI, *Auto a guida autonoma: profili di responsabilità civile e prospettiva di riforma*, in *Riv. dir. trasp.*, 2019, p. 351 ss.; M.M. COMENALE PINTO - E.G. ROSAFIO, *Responsabilità civile per la circolazione degli autoveicoli a conduzione autonoma. Dal grande fratello al grande conducente*, in *Riv. dir. trasp.*, 2019, p. 367 ss.

2. Dal conducente al robot. La categoria dei *Self Driving Vehicles*.

La dimensione sopra indicata, però, deve essere valutata con lo stato di avanzamento dell'automazione nel settore automobilistico.

Qui, appare opportuno, precisare che nel panorama dell'implementazione dell'intelligenza artificiale sui veicoli che circolano su strada, si assiste ad una scala graduale di sei livelli di automazione¹⁰. Il primo livello (livello zero) coincide con l'attuale sistema generalmente ancora in uso di un veicolo non automatizzato e completamente manovrato e controllato dal conducente, mentre l'ultimo livello (livello cinque) corrisponde all'ipotesi di un veicolo completamente automatizzato, senza conducente a bordo (*driverless car*)¹¹. I livelli intermedi si riferiscono ai diversi sistemi di assistenza alla guida: sistemi di *alert* relativi all'assistenza nelle manovre, passando attraverso i sistemi di *cruise control* per ridurre la velocità, per mantenere la corsia di marcia (già installati, invero, in molte autovetture), fino al livello quattro che è quello in cui il conducente guida il veicolo, oramai quasi totalmente automatizzato, per brevi tratti ed in determinate circostanze (guida autonoma piena o *full self-driving automation*).

Emerge, in definitiva, dal sistema di automazione sopra riportato, come il profilo più controverso e delicato sia quello dell'ultimo livello (sul quale difatti si sofferma il citato d. m. 28 febbraio 2018), mentre nelle altre fasi il conducente mantiene la manovra e la guida del veicolo, potendosi in parte dubitare della circostanza di cui al livello quattro in cui il conducente manterrebbe la guida per brevi tratti, alternandosi con i dispositivi automatici di controllo. In tale fattispecie, il conducente dovrebbe, comunque, mantenere un comportamento volto alla prudenza ed alla diligenza richieste al fine di riprendere il controllo del veicolo in quelle circostanze in cui gli strumenti tecnologici

¹⁰ La classificazione dei livelli di automazione è stata adottata dal Dipartimento dei Trasporti Usa ed elaborata dal SAE *International*, ente di normazione statunitense nel settore aerospaziale ed automobilistico.

¹¹ Per "veicolo a guida automatica" si intende "un veicolo dotato di tecnologie capaci di adottare e attuare comportamenti di guida senza l'intervento del guidatore, in determinati ambiti stradali e condizioni esterne" (art. 1, lett. f), d.m. 28 febbraio 2018).

restituiscano al conducente il controllo del mezzo o sia lo stesso conducente a rendersi conto del pericolo imminente o in altre particolari situazioni (avverse condizioni meteorologiche), per garantire l'incolumità delle persone.

Prova ne sia, a conforto di quanto appena sopra teorizzato, la disposizione dell'art. 1, lett. f), d.m. 28 febbraio 2018, che esclude dalla nozione di "veicolo a guida automatica" un "veicolo omologato per la circolazione delle strade pubbliche italiane secondo le regole vigenti e dotato di uno o più sistemi di assistenza alla guida, che vengono attivati da un guidatore al solo scopo di attuare comportamenti di guida da egli stesso decisi e che comunque necessitano di una continua partecipazione attiva da parte del conducente alla attività di guida".

Prima di analizzare le questioni giuridiche sopra indicate è necessario partire da una considerazione che assurge ad elemento imprescindibile nella struttura della circolazione stradale e della responsabilità civile automobilistica, il cui impianto normativo, oggi vigente in Italia, pone il conducente in una posizione centrale, quale figura assorbente i diversi criteri di imputazione della responsabilità civile *ex art. 2054 c.c.*

Appare qui significativo precisare, che per conducente di veicoli l'orientamento giurisprudenziale ritiene "colui che ha la direzione e la responsabilità della guida di un veicolo e si trovi quindi nella possibilità di compiere tutte quelle manovre che sono necessarie per la guida"¹², e cioè si tratta di un soggetto che ha l'effettiva disponibilità di quei congegni meccanici per la manovra del veicolo. In tal senso, si pone anche l'art. 8, legge 5 luglio 1995, n. 308 di ratifica ed esecuzione della Convenzione di Vienna dell'8 novembre 1969 sulla circolazione stradale, che prevede come ogni veicolo in movimento deve avere un conducente e ogni conducente deve possedere le qualità fisiche e psichiche necessarie ad essere nello stato fisico e mentale per condurre il veicolo. L'ultimo comma precisa che ogni conducente deve avere costantemente il controllo del proprio veicolo.

Al cospetto dello scenario, sopra indicato, di una progressiva introduzione di veicoli autonomi (e quindi definitivamente

¹² Cass. 7 novembre 1976, n. 4568, in *Resp. civ. prev.*, 1977, p. 316. Sulla nozione di conducente, cfr. S. SICA, *La circolazione dei veicoli*, Bologna, 2004, p. 24 ss; G. CASSANO (a cura di), *Trattato di infortunistica stradale*, Santarcangelo di Romagna, 2009, p. 75 ss.

senza la presenza del conducente), appare evidente il superamento sia dell'impianto normativo della responsabilità civile automobilistica, elaborato dall'art. 2054 c.c., connesso al settore normativo dell'assicurazione obbligatoria in materia, sia di quello di cui al Titolo V del codice della strada (d. lgs. 30 aprile 1992, n. 285) che delinea un pacchetto di norme comportamentali che il conducente di veicoli stradali è tenuto ad osservare per evitare criteri di imputazione di responsabilità civile¹³, nonché quelli di natura penale, allorquando si configurino i reati di lesioni stradali e di omicidio stradale¹⁴ e di natura sanzionatoria amministrativa¹⁵.

3. Il crepuscolo dell'art. 2054 c.c. Verso una nuova visione della responsabilità civile automobilistica.

Il quadro normativo sulla responsabilità civile per danni da circolazione di veicoli senza guida di rotaie è scolpito dall'art. 2054 c.c., norma ampiamente sondata dalla dottrina¹⁶, il quale pone una speciale responsabilità in via diretta in capo al condu-

¹³ Cfr. G. CASSANO (a cura di), *Diritto delle assicurazioni*, 2017, Milano, p. 367 ss.; A. LA TORRE (a cura di), *Le assicurazioni*, Milano, 2019, p. 957 ss. Sulla circolazione stradale, cfr. G.C. ANGELONI - E. BONASI BENUCCI, voce *Circolazione stradale*, in *Enc. Dir.*, VII, Milano, 1960, p. 21 ss. e M. DE GIORGI (a cura di), *Manuale pratico di infortunistica stradale*, Santarcangelo di Romagna, 2017.

¹⁴ Cfr. S. POLLASTRELLI - R. ACQUAROLI (a cura di), *Il reato di omicidio stradale*, 2017, Milano.

¹⁵ Cfr. A. CAGNAZZO - S. TOSCHEI - C. POZZI (a cura di), *Le sanzioni in materia di trasporto marittimo, aereo, terrestre e codice della strada*, Torino, 2012, p. 735 ss.

¹⁶ Sull'art. 2054 c.c., la letteratura giuridica è molto vasta. Per un'essenziale esposizione, cfr. D. PERETTI GRIVA, *La responsabilità civile nella circolazione*, Torino, 1959; A. LAGOSTENA BASSI - L. RUBINI, *La r.c. per la circolazione dei veicoli*, Milano, 1972; M. FRANZONI, *Fatti illeciti*, in *Comm. cod. civ.* Scialoja-Branca, Bologna-Roma, 1993, p. 642 ss.; F. PECCENINI, *La r.c. per la circolazione dei veicoli*, in *La responsabilità civile*, a cura di P. CENDON, XII, Torino, 1988, p. 3 ss.; S. SICA, *La circolazione dei veicoli*, cit., p. 23 ss.; C. CASTRONOVO, *La nuova responsabilità civile*, Milano, 2006, p. 430 ss.; G. CASSANO (a cura di), *Responsabilità ed assicurazione nella circolazione stradale*, Padova, 2007, p. 203 ss.; C.M. BIANCA, *Diritto civile, La responsabilità*, V, Milano, 2012, p. 754 ss.; G. BUFFONE (a cura di), *Responsabilità civile automobilistica*, Vicenza, 2016; AA. VV., *Responsabilità da circolazione stradale*, 2018, Milano; M. DE GIORGI, *Responsabilità civile automobilistica*, Roma, 2018.

cente, quale soggetto che materialmente cagiona il danno, ed in via indiretta ed in solido anche ad altri soggetti.

La disciplina dell'articolo in commento si snoda attraverso l'analisi del primo comma che pone in capo al conducente l'obbligo di risarcire il danno prodotto a persone o cose dalla circolazione del veicolo, se non prova di aver fatto tutto il possibile per evitare il danno. Il conducente, quindi, si esonera provando di avere adottato un comportamento prudente e diligente impiegando la perizia richiesta nella guida del veicolo e rispettando le norme comportamentali del codice della strada. L'orientamento della giurisprudenza è nel senso di consentire al conducente di avvalersi della prova liberatoria dimostrando l'altrui imprudenza, o una causa esterna inevitabile ed imprevedibile (caso fortuito, forza maggiore, comportamento della vittima, cattiva manutenzione della strada, ...) ¹⁷.

Il secondo comma prevede, che nel caso di scontro tra veicoli, si presume fino a prova contraria che ciascuno dei conducenti abbia concorso ugualmente a produrre il danno subito dai singoli veicoli.

Anche in tal caso, la figura del conducente è collocata al centro della presunzione di responsabilità. Ai fini applicativi del secondo comma occorre necessariamente la verifica di uno scontro tra veicoli (di una collisione) ¹⁸ che fa scattare la presunzione *iuris tantum* che può essere vinta dimostrando che il fatto dannoso è stato causato dalla condotta esclusiva dell'altro conducente. La giurisprudenza dell'Alta Corte ritiene che la presunzione di uguale responsabilità posta dalla norma abbia una funzione sussidiaria operando solo nel caso in cui non sia possibile accertare in concreto in quale misura la condotta dei

¹⁷ Sul punto, si vedano Cass. civ. 26.10.2017, n. 25421, in *Dir. giust.*, 2017, p. 1 ss., con nota di M. MAROTTA, *Nella circolazione dei veicoli in quanto attività pericolosa è configurabile il caso fortuito*; Cass. civ., 13.3.2012, n. 3966, in *Mass. giust. civ.*, 2012, p. 330; Cass. civ., 6.6.2006, n. 13268, in *Resp. civ. prev.*, 2007, p. 1467; Cass. civ. 11.3.2004, n. 4993, in *Foro it.*, 2004, I, c. 2108; Cass. civ. 19.7.2002, n. 10571, in *Danno e resp.*, 2003, p. 103 ss. In dottrina, cfr. A. LA SPINA, *Il sistema risarcitorio del danno da sinistro stradale: la tutela garantita e problemi ancora irrisolti*, in *Juscivile.it.*, 2015, p. 325 ss.

¹⁸ La giurisprudenza ritiene che il principio della presunzione di corresponsabilità si applica solo in caso di scontro tra veicoli e non nei casi in cui manchi la collisione (Cass. civ. 24.5.2006, n. 12370, in *Arch. giur. circol. e sinistri*, 2007, p. 35; Cass. civ. 23.7.2002, n. 10751, in *Mass. giust. civ.*, 2002, p. 1322).

due conducenti, coinvolti nello scontro, abbia cagionato l'evento dannoso, oppure quando le modalità del sinistro rimangano incerte¹⁹. Quindi, in difetto di prova liberatoria, totale o parziale, il concorso di ciascun conducente si presume sia eguale nella produzione del danno.

Il descritto sistema normativo dei primi due commi è imperniato esclusivamente sulla responsabilità della figura del conducente.

È nel terzo comma dell'art. 2054 c.c. che il legislatore prevede la responsabilità di altri soggetti.

Il proprietario del veicolo (o in sua vece l'usufruttuario o l'acquirente con patto di riservato dominio o, in base all'art. 91 codice della strada, il locatario del veicolo concesso in *leasing*) è responsabile in solido con il conducente, se non prova che la circolazione del veicolo sia avvenuta contro la sua volontà. Qui è chiara la *ratio* condivisibile della norma che assolve alla funzione di maggiormente tutelare le vittime degli incidenti stradali allargando la responsabilità a più soggetti, legati dal vincolo della solidarietà. Il proprietario, quindi, risponde indirettamente per un illecito commesso dal conducente sulla base di un'imputazione legale e potrà liberarsi dalla responsabilità se prova sia che il sinistro non è imputabile al conducente sia che la circolazione del veicolo sia avvenuta contro la sua volontà e cioè *prohibente domino*, non essendo sufficiente la circolazione *invito domino* (all'insaputa del proprietario), da accertarsi caso per caso e rimessa alla valutazione giudiziaria²⁰.

Plasmando quanto sopra normativamente tratteggiato sul versante della guida pienamente automatizzata, rimarrebbe il proprietario, quale unico soggetto responsabile per un evento dannoso non più commesso dal conducente (figura che uscirebbe completamente dai criteri di imputazione di responsabilità), il quale non avrebbe nemmeno la possibilità di esonerarsi provando che il sinistro non sia dipeso da una condotta di guida imprudente del conducente, né sarebbe in grado di incidere sulle dina-

¹⁹ In tal senso, Cass. civ. 18.7.2013, n. 17568, in *Dir. giust.*, 2013, p. 925 ss.; Cass. civ. 23.5.2013, n. 12667, in *Dir. giust.*, 2013, p. 850 ss.

²⁰ Cfr. Cass. civ. 27.6.2013, n. 16217, in *Dir. giust.*, 2013, p. 623 ss.; Cass. civ. 12.4.1990, n. 3138, in *Mass. giust. civ.*, 1990; Cass. civ. 1.8.2000, n. 10027, in *Danno e resp.*, 2001, p. 1062 ss. Cfr. G. TERRANOVA, voce *Responsabilità da circolazione di veicoli*, in *Dig. It., disc. priv., sez. civ.*, XVII, Torino, 1998, p. 89 ss.

niche automobilistiche. Il proprietario del veicolo, quindi, rimane coinvolto nella decisione di individuazione della responsabilità civile automobilistica, sia pure ripensando necessariamente ad una nuova struttura assicurativa.

Nello scenario normativo europeo, si rinvencono, in maniera significativa, due interventi legislativi in materia di responsabilità civile automobilistica automatizzata (*automated vehicles*).

Il Regno Unito si è dotato, nel 2018, di una legge “*Automated and Electric Vehicles Act*”, (preceduta da una proposta di legge denominata “*Vehicle Technology and Aviation Bill*”) che prevede la responsabilità in capo esclusivamente all’assicuratore nel caso di sinistro provocato da un veicolo automatizzato. Se, invece, il veicolo che circola non è assicurato, la responsabilità è in capo al proprietario del veicolo. L’assicuratore potrà esonerarsi dalla responsabilità se il sinistro sia direttamente riconducibile ad una indebita alterazione del sistema operativo installato sul veicolo da parte dell’assicurato o della quale questi fosse a conoscenza oppure in caso di mancato aggiornamento del *software*. Il legislatore inglese, dunque, mantiene una forma di responsabilità in capo al proprietario del veicolo che deve assicurarsi e che non deve compiere azioni volte ad ostacolare il corretto funzionamento dell’apparato tecnologico del veicolo, ma al tempo stesso, non sfugge la constatazione che non v’è alcuna forma di coinvolgimento nella responsabilità civile automobilistica inglese del soggetto che predispone il *software*, e cioè il soggetto produttore.

La Germania (primo Stato europeo a dotarsi di una legge sul traffico stradale) ha approvato definitivamente nel 2017 una modifica al codice della strada introducendo un’apposita disciplina sulla guida automatizzata. In sostanza, in caso di incidente stradale provocato da un errore del *software*, il soggetto responsabile tenuto al risarcimento dei danni è il proprietario del veicolo, a meno che non riesca a provare che il sinistro sia stato causato da forza maggiore. Alcuni aspetti della legge tedesca sono interessanti: v’è un profilo in materia di *privacy* relativa alla raccolta di dati da parte del *software* di guida ed è prevista l’attivazione e disattivazione manuale del sistema di guida automatizzata del veicolo quando a bordo c’è il conducente, il quale deve dialogare con il *software* di bordo in particolari circostanze (nei casi in cui sia il sistema di guida a richiedere al conducente di assumere il controllo e la manovra della guida). Sotto tale ultimo aspetto, sembrerebbe che il legislatore tedesco abbia voluto consentire la

circolazione delle *driverless cars* con un sistema misto in cui si preveda ancora la presenza a bordo del conducente, consentendo in un lasso di tempo di due anni di adeguare la normativa all'evoluzione della tecnologia²¹.

L'esame delle due discipline inglesi e tedesche mettono in luce come nei casi di sinistri in cui rimangono coinvolte autovetture completamente automatizzate l'unico soggetto responsabile sia il proprietario del veicolo anche quando si dimostri che il sinistro sia stato causato da un difetto o da un vizio del *software* installato, senza alcun coinvolgimento della casa costruttrice automobilistica (*car maker*).

Tale assunto consente di riprendere l'esame dell'ultimo comma dell'art. 2054 c.c. in base al quale le persone indicate nei commi precedenti (conducente, proprietario o l'usufruttuario o l'acquirente con patto di riservato dominio) sono responsabili, in ogni caso, dei danni derivati da vizi di costruzione o difetto di manutenzione del veicolo.

Si tratta di una forma di responsabilità oggettiva in cui i soggetti sopra indicati, e l'elenco è da intendersi in modo tassativo, rispondono in ogni caso per i danni da vizi di costruzione o da difetto di manutenzione del veicolo, indipendentemente dal loro comportamento colposo, fino al limite della forza maggiore, del caso fortuito o da causa diversa.

Non si ravvisa, nell'elenco dei soggetti normativamente indicati, la figura del produttore del veicolo nelle fattispecie dannose per vizi di costruzione e/o da difetto di manutenzione.

La fattispecie normativa dell'ultimo comma dell'art. 2054 c.c. è finita sotto la torsione della giurisprudenza che, in una fase anteriore al D.P.R. 24 maggio 1998, n. 224 (in materia di responsabilità per danno da prodotti difettosi, a sua volta poi modificato dal d.lgs. 2 febbraio 2001, n. 25, di recepimento della Direttiva comunitaria n. 85/374/CEE del 25 luglio 1985, ed infine abrogato dall'art. 146, d.lgs. 6 settembre 2005, n. 206 - codice del consumo), affermava la responsabilità del produttore per i

²¹ Cfr. M.S. LOSANO, *Il progetto di legge tedesco sull'auto a guida automatizzata*, in *Diritto dell'informazione e dell'informatica*, 2017, p. 1 ss. Appare opportuno precisare l'importanza del profilo della legge tedesca in materia di privacy che è un tema fortemente connesso all'intelligenza artificiale. Sul punto, cfr. S. CRISCI, *Intelligenza artificiale ed etica dell'algoritmo*, cit., p. 1796.

vizi di costruzione del veicolo al fine di garantire una maggiore tutela ai soggetti danneggiati, rompendo il principio tassativo dei soggetti ritenuti responsabili dall'ultimo comma dell'art. 2054 c.c. e superando la precedente impostazione che richiedeva il coinvolgimento eventualmente del produttore attraverso l'applicazione dell'art. 2043 c.c.²². Le prime applicazioni giurisprudenziali in materia di responsabilità del produttore si sono avute nei casi "Nissan" ed "Aprilia". Il primo caso riguardava un guasto nell'impianto frenante che comportava il bloccaggio delle ruote dell'autovettura mentre il secondo caso si riferiva al cedimento improvviso dello sterzo di un motociclo. Il Tribunale di Roma affermava che l'evento fosse riconducibile, con elevatissima probabilità, ad un difetto meccanico o di costruzione e quindi imputabile al solo produttore. Le cause ignote rimangono imputabili al produttore e non al consumatore. Si nota come in caso di mancanza di una prova piena della sussistenza del vizio, il produttore è stato ritenuto responsabile sulla base di un giudizio di elevatissima probabilità del difetto del prodotto²³.

Le due sentenze aprono uno scenario di imputazione di responsabilità del produttore di autovetture (case automobilistiche) in cui la correlazione tra prodotto sicuro e tutela del consumatore assume un punto centrale nello sviluppo giurisprudenziale e dottrinario successivo nel settore della circolazione stradale, imputando all'impresa il rischio del danno prodotto.

In questa prospettiva, la Suprema Corte ritiene che la responsabilità oggettiva del conducente o del proprietario non esclude la responsabilità del costruttore²⁴.

Oggi, la normativa sulla responsabilità del produttore per danni da prodotto difettoso è racchiusa negli articoli 114-127 del codice del consumo (d.lgs. 6.9.2005, n. 206), applicabile, quindi, anche al settore della circolazione stradale nei casi di danni che scaturiscono dal difetto di costruzione e/o dal difetto di manutenzione, rompendo la rigida impostazione di cui all'ultimo comma dell'art. 2054 c.c.

Al cospetto di simile scenario, appare opportuno analizzare

²² Si veda, G. CASSANO (a cura di), *Responsabilità ed assicurazione nella circolazione stradale*, cit., p. 243 ss.

²³ Trib. Roma 4,12,2003, in *Danno e resp.*, 2004, p. 527 ss.

²⁴ Cass. civ. 9.3.2004, n. 4754, in *Mass. giust. civ.*, 2004, p. 3. In materia, cfr. G. ALPA (a cura di), *La responsabilità del produttore*, 2019, Milano.

i punti salienti della normativa prevista nel codice del consumo²⁵.

L'art. 114 c. cons. apre con un principio generale in base al quale il produttore è responsabile del danno cagionato da difetti del suo prodotto²⁶. Non si tratta di una responsabilità oggettiva assoluta in quanto l'art. 118 c. cons. prevede un elenco di sei ipotesi in cui il produttore può esonerarsi dalla responsabilità aquiliana a suo carico. Di queste sei fattispecie esimenti, l'aspetto più delicato e controverso, e che ha riflessi nel settore automobilistico automatizzato, sembrerebbe essere costituito da quello di cui alla lett. e), in base al quale "se lo stato delle conoscenze scientifiche e tecniche, al momento in cui il produttore ha messo in circolazione il prodotto, non permetteva ancora di considerare il prodotto come difettoso".

Si tratta del c.d. "rischio di sviluppo" che ha avuto un'interpretazione da parte della Corte di Giustizia (con sentenza 29 maggio 1997, n. 300), sotto il profilo dell'onere probatorio in capo al produttore volto all'esonero della propria responsabilità, nel senso che "il produttore deve dimostrare che lo stato oggettivo delle conoscenze tecniche e scientifiche, ivi compreso il loro livello più avanzato, al momento della messa in commercio del prodotto considerato, non consentiva di scoprire il difetto di quest'ultimo"²⁷.

Nel campo dell'intelligenza artificiale non apparirebbe una facile operazione quella di stabilire quale sia lo stato delle conoscenze tecniche e scientifiche esistenti nel momento della messa in commercio del veicolo automatizzato, in un settore in continua e rapida evoluzione e caratterizzato da una complessa e sofisticata tecnologia.

Non meno preoccupante, è l'art. 120 c. cons. che prevede un onere probatorio gravoso a carico del consumatore-danneggiato (acquirente dell'autovettura), il quale deve dimostrare il difetto,

²⁵ Cfr. G. ALPA - L. ROSSI CARLEO (a cura di), *Commentario al codice del consumo*, Napoli, 2005; E. CESARO (a cura di), *I contratti del consumatore. Commentario al codice del consumo*, Padova, 2007; G. DE CRISTOFARO - A. ZACCARIA, *Commentario breve al diritto dei consumatori*, Padova, 2013; L. ROSSI CARLEO (a cura di), *Il diritto dei consumi*, Torino, 2015; V. CUFFARO (a cura di), *Codice del consumo*, Milano, 2019.

²⁶ Cfr. U. CARNEVALI, *Il difetto di progettazione negli autoveicoli* (nota a sentenza Trib. Pisa 16.3.2011), in *Resp. civ. prev.*, 2011, p. 2108 ss.

²⁷ In *Dir. comunitario e scambi internaz.*, 1998, p. 361 ss.

il danno subito e la connessione causale tra difetto e danno. Mentre, il produttore deve provare i fatti che escludono la sua responsabilità e su tale versante l'art. 116, primo comma, lett. b), consente al produttore di dimostrare che il difetto non esisteva ancora al momento in cui prodotto è stato messo in circolazione.²⁸ In tale prospettiva, nei veicoli automatizzati potrebbe risultare difficile provare il vizio o il difetto dell'algoritmo che controlla la manovra del veicolo o vizi nell'applicazione operativa dell'intelligenza artificiale.

Sui difetti di manutenzione, l'art. 2054, ultimo comma, c.c., mantiene fermo il principio applicativo di un dovere di diligenza del conducente e del proprietario nella manutenzione del veicolo. Ciò nondimeno, le manutenzioni sono oramai assolute dalle case automobilistiche e/o dai concessionari ai quali va imputata la responsabilità per vizi di manutenzione, causa dell'evento dannoso.

4. Considerazioni conclusive.

Le riflessioni superiormente svolte, mettono in luce un'ineguaglianza dell'art. 2054 c.c. a regolamentare la responsabilità civile automobilistica nel settore delle *driverless cars*.

Invero, l'art. 2054, ultimo comma, c.c., non è più in sintonia con l'ordinamento giuridico facendo gravare la responsabilità a soggetti che non hanno alcun coinvolgimento con i vizi di costruzione, semmai con i danni da circolazione stradale, per cui il legislatore, già sotto la vigenza del citato D.P.R. 24 maggio 1998, n. 224 e, successivamente, del codice del consumo, avrebbe dovuto coordinare la norma in esame con la disciplina per danni da prodotto difettoso.

La fuoriuscita della figura del conducente sembrerebbe lasciare il campo all'individuazione del proprietario dell'autovettura autonoma e del produttore (e cioè delle case automobilistiche) cui imputare i profili di responsabilità all'interno di una cornice normativa che, come sopra analizzata, non sembra essere però del tutto favorevole alle posizioni dei soggetti danneggiati e, in parte, penalizzante per lo stesso proprietario.

In tale prospettiva, inoltre, è da prendere in considerazione

²⁸ Cass. civ. 19.2.2016, n. 3258, in *Guida al diritto*, 2016, p. 51.

che i danni possono provenire non solo da difetti di costruzione o di produzione (tipici delle case automobilistiche), ma, anche, da difetti di progettazione del *software* di bordo e di impostazione degli algoritmi di base (produttore del *software*) ed, altresì, da difetti di comunicazione di dati tra le *driverless cars* e le *smart road* in un sistema tecnologico di correlazione e di interoperabilità (produttore o fornitore del sistema di comunicazione e/o gestore dell'infrastruttura intelligente e digitale), che non apparirebbero propriamente riconducibili alla figura del costruttore, né a quella del proprietario del veicolo connesso ed autonomo.

Ne consegue, quindi, che l'avviata fase di sperimentazione volta alla graduale sostituzione dei veicoli tradizionali con quelli autonomi e semi-autonomi (ed uno degli aspetti critici che si profila sarà proprio quello in cui, per un periodo di tempo, ad oggi, indefinibile, le auto a guida autonoma circoleranno insieme alle auto condotte dall'uomo), comporta necessariamente un profondo ripensamento della struttura della responsabilità civile automobilistica (tenendo conto delle diverse e nuove figure soggettive coinvolte), con riflessi immediati nel settore dell'assicurazione obbligatoria (a tutela dei soggetti danneggiati dalla circolazione stradale) ed in quello proprio della circolazione stradale (in cui si dovrà necessariamente rivedere il codice della strada).

LA DÉCISION EN DROIT ET CONTENTIEUX ADMINISTRATIFS FRANÇAIS

Pierre Serrand

SOMMAIRE: 1. Introduction. – 2. La décision en contentieux administratif: l'acte administratif décisoire. – A. La notion. – B. Les distinctions. – 3. La décision en droit administratif: l'acte administratif discrétionnaire. – A. Décision et autorité administrative. – B. Décision et juridiction administrative.

Introduction.

Les systèmes juridiques s'intéressent aux normes. Ils ne s'intéressent guère aux décisions. L'ordre juridique est en effet toujours présenté comme un agencement de normes ou de règles. Cet agencement est spécifique dans la mesure où ces normes ou règles apparaissent hiérarchiquement subordonnées les unes par rapport aux autres. Cette hiérarchie procède du contrôle juridictionnel. Le contrôle de constitutionnalité des lois permet ainsi de situer les normes constitutionnelles au-dessus des normes législatives. Le contrôle de conventionnalité des lois permet de donner son plein effet à l'article 55 de la Constitution: les engagements internationaux «ont une autorité supérieure à celle des lois». Le contrôle de légalité des actes administratifs permet de soumettre ces actes aux normes législatives. Cette hiérarchie normative est objective dans la mesure où elle s'impose de la même manière à tous les utilisateurs du droit. En effet, pour toutes les institutions françaises et pour tous les citoyens français, l'ordre juridique national comporte au sommet la Constitution; au-dessous les engagements internationaux; encore au-dessous les lois; et puis tout en bas les actes administratifs - réglementaires d'abord, individuels ensuite.

Cette composante normative concerne toutes les branches du droit, même le droit constitutionnel, ce qui n'est pas

évident. C'est ainsi que, dans la logique normativiste de Kelsen, la Constitution est la norme suprême, la norme des normes. Comment articuler cette notion de Constitution avec la notion d'ordre juridique, dont la caractéristique fondamentale est d'être un édifice à plusieurs étages superposés ? En effet, selon la théorie de la réalisation du droit par degré, un ordre juridique est une pyramide dans laquelle chaque norme, à la fois, sert de fondement à des normes inférieures et tire son fondement d'une norme supérieure. Dans la mesure où la Constitution est une norme parmi d'autres, quel en est le fondement ? On sait que Kelsen adopte une « hypothèse logico-transcendantale » selon laquelle la Constitution tire son fondement d'une norme qui n'est pas posée, mais supposée, et dont le contenu serait le suivant : « on doit se conduire conformément à la Constitution effectivement posée et efficace ». Cette manière de voir est contestable. N'est-il pas plus convaincant de soutenir, comme le font d'autres auteurs, que la Constitution est une décision du pouvoir constituant s'exprimant par le vote, soit d'une assemblée constituante, soit des citoyens par référendum ? La Constitution française du 4 octobre 1958 a ainsi été adoptée par le référendum du 28 septembre 1958 avec une majorité de 79,25% des voix. N'est-elle pas, dès lors, le produit d'une décision du peuple souverain ? Il semble ainsi plus pertinent de faire reposer l'ordre juridique sur une décision existante et posée par le souverain plutôt que sur une norme hypothétique et supposée. La place de la décision en droit constitutionnel n'apparaît pas seulement dans la volonté constituante, elle se manifeste aussi dans l'application de la Constitution. Les règles constitutionnelles sont en effet très majoritairement des normes d'habilitation qui répartissent des compétences entre les pouvoirs publics constitutionnels¹. Les actes pris sur le fondement de ces normes apparaissent très fréquemment comme des décisions qui expriment la libre volonté de leurs auteurs. Il suffit pour s'en convaincre de penser aux décisions du Président de la République de dissoudre l'Assemblée nationale, de recourir aux pleins pouvoirs de l'article 16, ou

¹ C'est en ce sens que, selon l'article 16 de la Déclaration des droits de l'homme et du citoyen de 1789, la Constitution est une séparation des pouvoirs.

d'envoyer les forces françaises sur une zone de guerre. La liberté du chef de l'État est d'autant plus grande que de tels actes sont, du point de vue du droit français, des actes de gouvernement injusticiables, faute de juge compétent pour les connaître et apprécier leur régularité juridique. Le Président est donc seul juge. De tels actes, qui expriment dès lors la libre volonté du président, sont ainsi de véritables décisions. La doctrine le confirme. Les constitutionnalistes, lorsque par exemple ils désignent le décret par lequel, en application de l'article 12 de la Constitution, le Président de la République prononce la dissolution de l'Assemblée Nationale, parlent toujours de la décision du chef de l'État, et non pas de la norme ou de la règle prise par lui.

Les décisions se manifestent non seulement en droit constitutionnel, elles apparaissent aussi en droit administratif. Les textes font d'abord de nombreuses références à cette notion. Pour se limiter au code des relations entre le public et l'administration (CRPA), ce terme apparaît à 186 reprises². Le code utilise ainsi les expressions suivantes:

- «décision(s) administrative(s)» (art. L. 134-2; art. L. 212-2; art. L. 311-2; art. L. 410-1; art. L. 411-2; art. L. 431-1),
- «décisions de l'administration» (art. L. 212-3),
- «décisions réglementaires» (art. L. 200-1),
- «décision(s) individuelle(s)» (art. L. 121-1 et 121-2; art. L. 200-1; art. L. 211-7; art. L. 231-4; L. 311-3-1),
- «décisions ni réglementaires ni individuelles» (art. L. 200-1; art. L. 221-7),
- «décisions administratives individuelles» (art. L. 211-3),
- «décision individuelle expresse» (art. L. 221-8),
- «décision implicite ou expresse» (art. R. 112-4),
- «décision de refus» (art. L. 311-14; art. R. 343-1; art. R. 343-3),
- «décision implicite de refus» (art. R. 343-5),
- «décision de rejet» (art. L. 231-4; art. L. 411-7),

² En outre, le mot «décision» apparaît dans plusieurs intitulés des titres du CRPA: «Décisions soumises au respect d'une procédure contradictoire préalable»; «L'association du public aux décisions prises par l'administration»; «Participation du public aux décisions locales»; «Décisions ni réglementaires ni individuelles»; «Décisions individuelles»; «Les décisions implicites».

- «décision implicite de rejet» (art. L. 114-3; art. L. 232-4),
- «décision d'acceptation» (art. L. 231-1; art. R. 323-5),
- «décision implicite d'acceptation» (art. L. 114-3; art. L. 232-3),
- «décision créatrice de droits» (art. L. 242-2, 3, 4 et 5; art. L. 411-4),
- «décision non créatrice de droits» (art. L. 411-4).
- «décision d'attribution» d'un droit (art. L. 114-5-1),
- «décision(s) défavorable(s)» (art. L. 322-5; L. 342-1),
- «décision écrite» (art. L. 311-14, art. L. 322-5),
- «décision motivée» (art. R. 323-6).

La jurisprudence utilise également très souvent le mot décision. Pour se limiter à la jurisprudence du Conseil d'État rendue en 2018, ce terme apparaît dans 2390 de ses arrêts ou ordonnances. La décision est donc partout. Elle intéresse désormais la doctrine puisque la thèse de Benjamin Defoort lui est justement consacrée³. Il convient, si ce n'est de la définir, tout au moins de préciser ce qu'elle peut signifier.

Le droit administratif nous semble utiliser le mot «décision» dans deux sens différents. Autrement dit, deux notions de décision peuvent être distinguées selon que l'on prend en considération les effets de la décision ou la situation juridique qui lui sert de fondement.

Si l'on regarde les effets de la décision, ce qu'elle produit; c'est-à-dire si on l'examine du point de vue de ses destinataires, la décision administrative est synonyme d'acte administratif décisoire. La décision est ici l'acte administratif dont les dispositions ont un caractère impératif, c'est-à-dire dont les dispositions s'imposent à ses destinataires. *Décider, c'est ici imposer.*

Si l'on regarde le fondement juridique de la décision, c'est-à-dire si l'on examine la décision du point de vue des normes juridiques qui lui servent de base légale, la décision administrative est synonyme d'acte administratif discrétionnaire. La décision est ici l'acte administratif exprimant un pouvoir discrétionnaire de l'autorité

³ B. DEFOORT, *La décision administrative*, Paris, 2015.

administrative, c'est-à-dire l'acte manifestant une liberté de choix. *Décider, c'est ici choisir.*

La première notion de décision est celle du droit du contentieux administratif. Elle ressort en effet des règles juridiques qui régissent la recevabilité du recours devant le juge administratif. La seconde notion de décision est celle du droit administratif général, tel que la doctrine en rend compte. Elle ressort en effet de ce que la doctrine analyse comme étant le pouvoir discrétionnaire de l'Administration. Reprenons ces deux points successivement.

2. La décision en contentieux administratif: l'acte administratif décisoire.

La notion de décision apparaît en droit du contentieux administratif. Elle est liée à la règle dite de la «décision préalable». Cette règle est ainsi formulée à l'article R. 421-1 du code de justice administrative: «La juridiction ne peut être saisie que par voie de recours formé contre une décision». Elle signifie qu'avant de saisir le juge administratif, il faut provoquer une décision de l'Administration, et le recours juridictionnel ne peut être dirigé que contre cette décision. Le recours contentieux n'est ainsi recevable que s'il est exercé contre un acte décisoire, c'est-à-dire une manifestation de volonté porteuse d'une décision. Il convient d'identifier cette notion avant de procéder à des distinctions.

A. La notion.

Comment reconnaît-on un acte décisoire ? L'acte décisoire se reconnaît à son contenu. C'est en effet ce que nous enseigne l'arrêt de Section du 18 décembre 2002, Mme Duvignères (n° 233618), dans lequel le Conseil d'État se prononce sur la nature juridique des circulaires et instructions ministérielles. Il considère que «les dispositions impératives à caractère général d'une circulaire ou d'une instruction doivent être regardées comme faisant grief». Le recours contentieux est par conséquent recevable contre un acte contenant de telles dispositions, cet acte étant décisoire. Depuis cet arrêt, il convient par conséquent de distinguer selon que les dispositions des actes à qualifier sont, soit impératives, soit

purement indicatives. Dans le premier cas, l'acte est décisoire; dans le second, il est non-décisoire. Comme le précise Pascale Fombeur dans ses conclusions sur l'arrêt Duvignères⁴: par dispositions impératives à caractère général, il faut entendre toutes les dispositions au moyen desquelles une autorité administrative vise, soit à créer des droits ou des obligations, c'est-à-dire à poser une norme nouvelle, soit à imposer une interprétation du droit applicable en vue de l'édiction de décisions, ce qui revient à déterminer la norme applicable. Le critère du caractère impératif ou non de l'instruction ou de la circulaire n'est pas toujours facile à mettre en œuvre. Il faut en effet distinguer les simples indications et les prescriptions auxquelles les destinataires de l'acte doivent se conformer. Pour cela, le juge analyse les termes, le style, le ton de la circulaire ou de l'instruction. Il recherche également l'intention de son auteur. Il examine encore - et peut-être surtout - la façon dont l'acte est perçu par ses destinataires. Le juge considère ainsi que l'instruction ou la circulaire est décisoire si les agents de l'Administration se sentent liés par elle, s'ils prennent soin de ne pas s'en écarter. L'analyse porte donc sur la force normative de l'acte, c'est-à-dire le modèle de conduite qu'il prescrit. Si la circulaire ou l'instruction attaquée contient des dispositions de caractère impératif et général, elle est regardée comme un acte décisoire susceptible de recours. Si ce n'est pas le cas, c'est-à-dire si l'auteur de l'acte a seulement indiqué à ses services comment appliquer le droit applicable, la requête est irrecevable car cet acte est considéré comme non décisoire.

Toujours selon l'arrêt Duvignères, les circulaires ou instructions constitutives d'une décision sont de deux types. Elles sont réglementaires lorsqu'elles «fixent, dans le silence des textes, une règle nouvelle». Elles sont interprétatives lorsqu'elles «prescrivent d'adopter» une interprétation d'un texte. En effet, selon le Conseil d'État, «l'interprétation que par voie, notamment, de circulaires ou d'instructions l'autorité administrative donne des lois et règlements qu'elle a pour mission de mettre en œuvre» est susceptible d'être déféré au juge de l'excès de pouvoir lorsque, ayant un caractère

⁴ P. FOMBEUR, *Revue française de droit administratif*, 2003, p. 280.

impératif, elle fait grief. L'interprétation d'un texte est donc constitutive d'une décision dès lors que cette interprétation s'impose à ses destinataires. Cette solution est intéressante dans la mesure où elle confirme la théorie dite «réaliste» de l'interprétation. On sait en effet qu'il est possible de concevoir l'interprétation d'un texte de deux manières. On peut d'abord estimer qu'il s'agit d'un acte de connaissance. C'est la conception traditionnelle. L'interprète doit, par une opération de la connaissance, déterminer le sens qui se trouve dans le texte à interpréter. Il peut pour cela utiliser une méthode d'interprétation: méthode téléologique, raisonnement *a contrario* ou *a fortiori*... Mais on peut également estimer qu'interpréter est un acte de volonté. Il s'agit pour l'interprète de choisir entre les différentes significations possibles d'un texte, et ce choix procède de la seule volonté de l'interprète. Or, dans la mesure où une règle de droit est une signification, et parce que l'interprète détermine cette signification, l'interprète devient l'auteur de la règle de droit. C'est déjà ce qu'affirme Hobbes dans le Chapitre XXVI du *Léviathan* :

«Ce n'est pas dans la lettre que réside la nature de la loi: c'est dans ce qu'elle contient, ce qu'elle veut dire: autrement dit, dans l'interprétation authentique de la loi, qui n'est autre que la pensée ou le sens du législateur. L'interprétation de toutes les lois dépend donc de l'autorité souveraine, et les interprètes ne peuvent être personne d'autre que ceux qu'aura institués le souverain, à qui seul les sujets doivent obéissance. Autrement, l'astuce d'un interprète peut faire exprimer à la loi une pensée ou un sens contraire à celle du souverain, et par ce moyen l'interprète devient législateur.»

Or, le choix fait par l'interprète du sens qu'il convient de donner au texte est une pure opération de volonté, c'est-à-dire une véritable décision. Le Conseil d'État nous semble donc avoir raison de regarder, comme il le fait dans l'arrêt *Du vignères*, l'acte par lequel une autorité administrative interprète un texte, dès lors que cette interprétation a un caractère impératif, comme un acte décisoire.

Au terme de l'analyse, la décision administrative est un acte juridique, c'est-à-dire la manifestation de volonté porteuse d'une norme juridique, dont les dispositions s'imposent à ses

destinataires. Ces décisions sont nombreuses. Il convient de procéder à des distinctions.

B. Les distinctions.

Trois distinctions peuvent être effectuées selon qu'est pris en compte: le contenu de la décision, sa forme, ou l'autorité qui la prend. La première distinction est matérielle; la deuxième est formelle; la troisième est organique.

1) Distinction matérielle

La distinction matérielle ressort de l'article L. 200-1 du CRPA: «Les actes administratifs unilatéraux décisifs comprennent les actes réglementaires, les actes individuels et les autres actes décisifs non réglementaires. Ils peuvent être également désignés sous le terme de décisions, ou selon le cas, sous les expressions de décisions réglementaires, de décisions individuelles et de décisions ni réglementaires ni individuelles.» Il ressort de la jurisprudence que le mot «décision» est plutôt utilisé pour désigner les actes individuels. Les décisions «ni réglementaires ni individuelles» sont ce que l'on appelle des «décisions d'espèce». Examinons-les successivement.

La décision réglementaire est souvent présentée comme ayant un «caractère général et impersonnel»⁵. Elle est impersonnelle en ce sens qu'elle édicte une norme ayant vocation à régir, non pas un individu nommément désigné, mais une collectivité abstraite de personnes. C'est donc le plus souvent un acte général, comme l'est par exemple la décision d'un maire d'instaurer dans sa commune une zone réservée aux piétons, laquelle s'applique à tous ceux qui circulent sur les voies communales. Mais il arrive qu'une décision réglementaire ne possède pas ce caractère général. C'est le cas, comme l'a jugé le Conseil d'État, du décret relatif au statut de l'administrateur de la Comédie-Française qui, en raison du caractère unique de cet emploi, n'a pas un caractère

⁵ CE, Sect., 1^{er} juillet 2016, Institut d'ostéopathie de Bordeaux, n° 393082.

général, mais qui, n'identifiant personne nommément, a bien ce caractère d'abstraction permettant de le regarder comme un «décret réglementaire»⁶.

Parmi les décisions réglementaires, une place à part doit être faite aux décisions relatives à l'organisation d'un service public. Le lien est ici suffisamment fort que certains arrêts font apparaître que c'est parce qu'une décision est relative à l'organisation d'un service public que cette décision est réglementaire. Le Conseil d'État considère par exemple qu'un arrêté déclarant vacant un office de greffier de tribunal de commerce et précisant les modalités de dépôt des candidatures à la succession «est relatif à l'organisation du service public de la justice et présente ainsi un caractère réglementaire»⁷. Dans le prolongement, pour dénier à un acte un caractère réglementaire, le Conseil d'État peut préciser que cet acte, «d'une part, est dépourvu de caractère général et impersonnel et, d'autre part, n'a pas, par lui-même, pour objet l'organisation d'un service public»⁸. Cette démarche conduit à adopter une double définition de la décision réglementaire. Est en effet réglementaire: d'une part la décision de caractère général et impersonnel; d'autre part la décision qui, sans être nécessairement générale, est relative à l'organisation d'un service public.

À la différence de la décision réglementaire, la décision individuelle s'adresse à une ou plusieurs personnes nommément désignées. Cet acte juridique contient donc une norme qui s'impose à une ou plusieurs personnes identifiées par leur nom. Il peut s'agir de personnes physiques ou morales, de personnes privées ou de personnes publiques. C'est le cas d'un décret portant nomination d'un fonctionnaire, d'un permis de construire, d'une autorisation de police. Le nombre des personnes visées est indifférent. La décision d'autoriser une liste de candidats à se présenter à un concours est ainsi, en l'absence de caractère abstrait, une décision individuelle.

Enfin, le législateur consacre l'existence de «décisions ni

⁶ CE, Ass., 13 juillet 1962, Bréart de Boisanger, dans *Les grandes décisions de la jurisprudence. Droit administratif*, 17^e éd., 2017, Paris.

⁷ CE, 11 février 2015, M. G... B..., n^o 367884.

⁸ CE, 19 juin 2017, SAGESS, n^o 403316.

réglementaires ni individuelles» (art. L. 200-1 CRPA). La jurisprudence en donne des exemples. C'est ainsi qu'une décision portant reclassement d'une section de route nationale dans la voirie d'une collectivité territoriale «ne constitue pas une décision réglementaire et ne présente pas davantage le caractère d'une décision administrative individuelle»⁹. Il s'agit d'une décision d'espèce, c'est-à-dire d'une décision qui édicte une norme qui n'est ni générale, ni individuelle, mais qui est une norme d'espèce, c'est-à-dire se rapportant à une situation ou à une opération particulière. En d'autres termes, il s'agit d'une décision appliquant à une espèce particulière une réglementation existante qui n'est pas modifiée. Sont par exemple des décisions d'espèce: la décision d'ouvrir un concours de recrutement dans la fonction publique¹⁰, la décision classant un bien monument historique¹¹, les déclarations d'utilité publique¹², ainsi que les décisions créant une zone d'aménagement concerté¹³.

2) Distinction formelle

La distinction formelle conduit à opposer les décisions expresses et les décisions implicites.

Les décisions expresses ont, le plus souvent, une forme écrite: décret, arrêté, circulaire, instruction... Elles comportent en principe des visas et un dispositif rédigé en articles. Mais les décisions expresses peuvent aussi avoir une autre forme. C'est ainsi qu'un message électronique émanant d'un directeur d'une Administration centrale et ayant un caractère impératif est ainsi un acte administratif décisoire¹⁴. En outre, certaines décisions administratives sont prises sur le fondement d'un algorithme, comme le sont par exemple les décisions individuelles prises dans le cadre de l'application Parcoursup. La forme de la décision administrative peut encore être verbale. Une décision administrative peut ainsi être révélée par

⁹ CE, 25 septembre 2009, Commune de Coulomby, n° 310873.

¹⁰ CE, 28 septembre 1988, Marciano, n° 92793.

¹¹ CE, 7 février 1992, Ministre de la Culture, n° 118488.

¹² CE, 11 février 1983, Commune de Guidel, n° 41233.

¹³ CE, 2 décembre 1977, Comité de défense de l'environnement de Mâcon-Nord, n° 00843.

¹⁴ CE, 3 février 2016, Ordre national des infirmiers, n° 381203.

un discours politique du Premier ministre¹⁵. La forme peut encore être gestuelle: une décision administrative émane à chaque fois de l'agent de police qui, à un carrefour, fait signe aux piétons de passer, puis aux voitures de circuler. Elle peut également être sonore: la sirène des pompiers exprime une décision donnant l'ordre aux secours d'intervenir. Certaines décisions administratives n'ont même aucune forme, comme le sont les décisions implicites.

Les décisions implicites résultent du silence que l'autorité administrative garde à une demande qui lui est faite. En principe, «le silence gardé pendant deux mois par l'Administration sur une demande vaut décision d'acceptation» (art. L. 231-1 CRPA). On applique ainsi l'adage: «Qui ne dit mot consent». Le champ d'application de cette règle est très large puisqu'elle s'applique à toutes les procédures prévues par un texte débouchant sur une décision individuelle. Elle s'applique en outre aux Administrations d'État, aux collectivités territoriales, aux établissements publics administratifs, et aux organismes et personnes de droit public et de droit privé chargés d'une mission de service public administratif, y compris les organismes de sécurité sociale (art. L. 100-3 CRPA). Elle ne s'applique en revanche pas aux relations que les Administrations entretiennent avec leurs agents; le silence des autorités administratives vaut donc ici refus (art. L. 231-4 CRPA).

À ce principe, il existe deux séries d'exception. D'abord, le délai qui fait naître la décision implicite peut ne pas être de deux mois, mais être, soit plus court, soit plus long. Le délai est ainsi de 10 jours pour l'autorisation d'une manifestation publique de boxe; à l'inverse, le délai est de trois mois pour l'inscription d'un enfant en école maternelle. Ensuite, le silence de l'Administration peut parfois valoir rejet de la demande. Cette exception qu'est la décision implicite de rejet est notamment susceptible de s'appliquer dans les cinq hypothèses prévues par l'article L. 231-4 du CRPA. Il en va ainsi, par exemple, lorsque «la demande présente un caractère financier». Les exceptions au principe sont suffisamment importantes pour

¹⁵ CE, 15 mars 2017, Association «Bail à part, tremplin pour le logement», n° 391654.

que, en pratique, on puisse considérer que le silence de l'Administration vaut le plus souvent rejet de la demande.

3) Distinction organique

La distinction organique conduit à classer les décisions administratives selon les autorités qui les prennent.

Il y a d'abord les décrets, qui sont édictés par le Président de la République et par le Premier ministre, selon que le décret est ou non délibéré en conseil des ministres.

Il y a ensuite les arrêtés, qui sont pris par toutes les autres autorités administratives individuelles: arrêté ministériel lorsque la décision est prise par un ministre, arrêté préfectoral lorsqu'elle l'est par un préfet, arrêté municipal lorsqu'elle l'est par un maire...

Enfin, il y a les délibérations, qui sont adoptées par les autorités administratives collégiales des collectivités territoriales et des établissements publics. Le conseil municipal adopte ainsi des délibérations (art. L. 2121-29 Code général des collectivités territoriales), de même que le conseil d'administration de l'université (art. L. 712-1 Code de l'éducation).

Voilà comment se présente, du point de vue de la procédure administrative contentieuse, la décision administrative. De ce point de vue, rappelons qu'elle est synonyme d'acte administratif décisoire, de telle manière que la norme juridique contenue dans la manifestation de volonté s'impose à ses destinataires puisque les dispositions de l'acte ont un caractère impératif. La décision administrative est alors étudiée en prenant en compte ses effets à l'égard de ses destinataires.

Mais il est également possible d'appréhender la décision administrative en examinant le fondement de cet acte juridique. Il s'agit de rechercher si l'autorité administrative dispose d'une liberté de choix pour déterminer le contenu de l'acte juridique qu'elle édicte. De ce point de vue, la décision administrative apparaît comme l'expression d'un pouvoir discrétionnaire de l'autorité administrative qui la prend.

3. La décision en droit administratif: l'acte administratif discrétionnaire.

La notion de décision apparaît en droit administratif

général, plus précisément dans la théorie du droit administratif. La décision suppose ici une liberté de choix. Il s'agit de choisir entre plusieurs possibilités. Ce choix est le fruit d'une volonté autonome, c'est-à-dire d'une volonté qui n'est pas déterminée par une norme qu'il conviendrait d'exécuter. En ce sens, *la décision s'oppose à l'exécution*. La décision administrative se manifeste ici dans les domaines où l'autorité administrative dispose d'une latitude d'action, d'une possibilité de choisir. On dit en droit administratif que l'autorité administrative est en situation de pouvoir discrétionnaire. C'est ce qu'il convient d'abord de préciser.

A. Décision et autorité administrative.

«Il y a pouvoir discrétionnaire toutes les fois qu'une autorité agit librement sans que la conduite à tenir lui soit dictée à l'avance par une règle de droit» (L. Michoud, 1913). Autrement dit, le pouvoir discrétionnaire est «un plein pouvoir d'appréciation pour celui qui en est investi, à condition qu'il l'exerce légalement et dans le but pour lequel il a été créé» (J. Helbronner, 1912). Sous l'inévitable variété des formulations, c'est l'idée de liberté, cette place faite à l'initiative, cette faculté d'appréciation, cette existence d'une possibilité de choix, qui caractérise le pouvoir discrétionnaire.

Le pouvoir discrétionnaire résulte d'une certaine situation de l'ordre juridique: l'absence de norme préétablie. Le droit laisse l'Administration libre d'agir. L'autorité administrative est alors libre de son choix, et ce choix exprime une véritable décision. Deux exemples tirés de la jurisprudence permettent de l'illustrer.

Le premier exemple concerne la nomination aux emplois dits «à la décision du Gouvernement». Ils sont ainsi qualifiés car leurs titulaires sont étroitement associés à la mise en œuvre de la politique gouvernementale. Ces emplois sont régis par l'article 25 de la loi du 11 janvier 1984 portant dispositions statutaires relatives à la fonction publique de l'État. Cet article renvoie au décret n° 85-779 du 24 juillet 1985, qui détermine pour chaque Administration et service «les emplois supérieurs pour lesquels les nominations sont laissées à la décision du Gouvernement». Ce décret précise en outre que ces nominations sont «essentiellement révocables». Sur le fondement de ces textes, le

préfet de la région Provence-Alpes-Côte-d'Azur fut nommé par décret chef du service de l'inspection générale de l'Administration. Le Syndicat autonome des inspecteurs généraux et inspecteurs de l'administration au ministère de l'intérieur attaqua ce décret devant le Conseil d'État. L'arrêt du 11 juillet 2012 (n° 348064), rendu par l'Assemblée du Conseil d'État, nous donne une première illustration de ce qu'est une décision prise dans le cadre d'un pouvoir discrétionnaire, car c'est évidemment le choix de l'autorité administrative qui était contesté. Comme le précise le rapporteur public sur cet arrêt, la caractéristique principale des emplois à la décision du Gouvernement «tient à ce que, par dérogation aux règles de la fonction publique, le gouvernement est libre d'en choisir le titulaire et de le révoquer à tout moment»¹⁶. La seule limite juridique qui pèse sur le Gouvernement réside dans les dispositions de l'article 6 de la Déclaration des droits de l'homme de 1789, selon lesquelles «tous les citoyens (...) sont également admissibles à toutes dignités, places et emplois public, selon leur capacité et sans autre distinction que celle de leurs vertus et de leurs talents». Le Conseil constitutionnel considère en effet qu'en vertu de ces dispositions, le choix du Gouvernement «doit être fait en prenant en considération les capacités requises pour l'exercice des attributions afférentes à l'emploi» (n° 2010-94 QPC). On le voit, le Gouvernement est en situation de pouvoir discrétionnaire pour nommer aux emplois en cause. Il est en effet très largement libre de nommer qui il veut. L'acte juridique portant nomination à un tel emploi procède moins des règles juridiques applicables que du libre choix de l'autorité administrative. En ce sens, cet acte administratif de nomination exprime bien une décision.

Le deuxième exemple concerne le choix dont dispose l'Administration, lorsqu'elle a besoin d'une prestation, d'internaliser ou d'externaliser cette prestation. Comme l'indique l'article L. 1 du tout récent Code de la commande publique¹⁷: «Les acheteurs et les autorités concédantes

¹⁶ N. ESCAUT, Conclusions sur CE, Ass., 11 juillet 2012, Syndicat autonome des inspecteurs généraux et inspecteurs de l'administration, dans *Revue française de droit administratif*, 2012, p. 954.

¹⁷ Cette codification a été réalisée par l'ordonnance du 26 novembre 2018, et le code est entré en vigueur le 1^{er} avril 2019.

choisissent librement, pour répondre à leurs besoins, d'utiliser leurs propres moyens ou d'avoir recours à un contrat de la commande publique». Cela signifie que, pour obtenir la réalisation d'une prestation de travaux, de fournitures ou de services, l'Administration a la possibilité, soit de réaliser cette prestation en interne en recourant à ses propres services, soit de recourir à un tiers en passant avec lui un contrat de la commande publique¹⁸. Les collectivités publiques disposent ici d'une latitude d'action, d'une liberté d'appréciation. Le choix de l'autorité administrative de confier ou non à un tiers la gestion d'une telle mission n'est donc pas déterminé par la règle de droit. Il l'est d'ailleurs souvent par des considérations politiques, certaines majorités politiques étant plus favorables à l'externalisation, alors que d'autres préfèrent la gestion directe. Ce choix procède donc de la seule volonté de l'autorité administrative. Il est dès lors moins l'expression d'une norme que d'une décision. Le Conseil d'État en déduit que «la décision par laquelle une collectivité fait le choix de déléguer, ou non, un service public ou une partie de celui-ci» «échappe au contrôle du juge de l'excès de pouvoir»¹⁹.

On pourrait penser à travers cet exemple que le pouvoir discrétionnaire de l'Administration conditionne le contrôle que le juge exerce sur l'acte exprimant ce pouvoir. L'impression est en effet que la décision qui exprime la liberté de choix dont dispose l'autorité administrative n'est pas soumise au contrôle juridictionnel car le juge souhaite restreindre son contrôle pour ne pas substituer sa propre décision à celle de l'Administration. L'étendue du contrôle juridictionnel procéderait ainsi de la situation plus ou moins discrétionnaire dans laquelle l'ordre juridique placerait l'Administration. Toutefois, comme nous avons tenté de le montrer ailleurs²⁰, la

¹⁸ Cette solution est ancienne pour les missions de service public: «lorsque des collectivités publiques sont responsables d'un service public, elles peuvent (...) décider de confier sa gestion à un tiers»; mais elles «peuvent aussi décider d'en assurer directement la gestion» (CE, Sect., 6 avril 2007, Commune d'Aix-en-Provence, n° 284736).

¹⁹ CE, 27 janvier 2011, Commune de Ramatuelle, n° 338285.

²⁰ P. SERRAND, *Le contrôle juridictionnel du pouvoir discrétionnaire de l'administration à travers la jurisprudence récente*, dans *Revue du droit public*, 2012, n° 4, p. 901; P. SERRAND, *Droit administratif*, tome 2: *Les obligations administratives*, 2^e éd., 2019, Paris, n° 76.

situation de l'ordre juridique qui donne naissance au pouvoir discrétionnaire doit être appréciée en aval du contrôle juridictionnel. En effet, si le texte laisse libre l'Administration, mais que le juge, en interprétant ce texte ou en lui ajoutant des conditions, détermine l'action de l'Administration, celle-ci n'est plus libre d'agir. La décision administrative procède alors d'une autre décision: celle du juge administratif. C'est ce second aspect que nous souhaiterions examiner maintenant.

B. Décision et juridiction administrative.

La situation de l'ordre juridique qui donne naissance au pouvoir discrétionnaire et, par conséquent, à la décision administrative, doit être appréciée en aval du contrôle juridictionnel. En effet, si le texte est silencieux et laisse libre l'Administration d'agir, le juge peut décider de pallier le silence du texte en imposant à l'Administration de respecter des règles afin que son pouvoir, certes discrétionnaire, ne soit pas pour autant arbitraire. Dans ce cas, le juge détermine l'action de l'Administration, et celle-ci n'est dès lors plus libre d'agir. Un exemple, tiré de l'arrêt du 20 juillet 1990, Préfet de la Savoie c. M. Cavdar (n° 117364), permet de l'illustrer. En l'espèce, M. Cavdar est resté en France alors que lui a été notifiée une décision de refus de séjour. Sa demande effectuée auprès de l'Office français de protection des réfugiés et apatrides a été rejetée. Il n'a pas fait de recours contre cette décision de refus. Il se trouve par conséquent, selon la législation alors applicable, dans la possibilité de faire l'objet d'une mesure de reconduite à la frontière. Le préfet prononce cette mesure. Saisi d'un recours contre cette décision, le Conseil d'État crée une condition juridique supplémentaire s'imposant à l'Administration. Après avoir précisé que, selon les textes, le requérant pouvait faire l'objet d'une mesure de reconduite à la frontière, il considère «toutefois qu'il appartient au préfet, même dans ce cas, de vérifier si la mesure de reconduite ne comporte pas de conséquences d'une gravité exceptionnelle sur la situation personnelle ou familiale de l'intéressé». Puis, appliquant cette règle nouvelle à l'espèce, il estime «que M. Cavdar a soutenu sans être contredit qu'à la date de la décision attaquée son épouse bénéficiait d'un titre de séjour délivré dans l'attente d'une décision sur sa demande d'admission au statut de réfugié,

qu'elle était enceinte et de santé fragile». La conséquence pour l'Administration est «qu'en ordonnant dans de telles conditions et à cette date la reconduite à la frontière de M. Cavdar le préfet de Savoie a commis une erreur manifeste dans l'appréciation des conséquences de ladite mesure». L'acte administratif est ici illégal parce qu'il est contraire à une règle de droit qui provient, non pas du texte, mais du juge. L'autorité administrative semblait en situation de pouvoir discrétionnaire à la lecture du texte. Elle ne l'est pas en raison de l'intervention du juge.

Il apparaît par conséquent que, si le pouvoir discrétionnaire est une liberté, cette liberté doit être appréciée en aval du contrôle juridictionnel, car le juge peut en effet transformer le pouvoir discrétionnaire en compétence liée. Et pour cela, le juge prend une véritable décision. La décision du juge, c'est-à-dire son choix «anormé», apparaît clairement ici dans la mesure où la volonté du juge se manifeste dans le silence de la loi. Il s'agit, pour reprendre les termes de Portalis, de pratiquer l'interprétation par voie de doctrine; elle «consiste à saisir le vrai sens des lois, à les appliquer avec discernement, et à les suppléer dans les cas qu'elles n'ont pas réglés»²¹. C'est ainsi que, pour reprendre encore les termes de Portalis: «Quand la loi est claire, il faut la suivre; quand elle est obscure, il faut en approfondir les dispositions.»²² On comprend que, pour reprendre l'arrêt Cavdar, la loi étant aux yeux du juge insuffisante ou obscure, le Conseil d'État ait pu, par un acte de volonté exprimant une décision, en approfondir les dispositions. Mais qu'en est-il lorsque la loi est claire? Quand la loi est claire, il faut la suivre nous dit Portalis. Il s'agit, pour le juge, d'appliquer la loi. Le juge doit être ici, pour reprendre les termes de Montesquieu, la bouche qui prononce les paroles de la loi. N'y-a-t-il alors aucun libre choix, aucune décision? Un autre arrêt du Conseil d'État permet de répondre à cette question: le grand arrêt Dame Lamotte du 17 février 1950²³.

²¹ J. E. M. PORTALIS, *Discours préliminaire sur le projet de Code civil*, prononcé le 1^{er} pluviôse an IX.

²² *Ibidem*.

²³ *Les grands arrêts de la jurisprudence administrative*, 21^o éd., 2017, Paris.

En l'espèce, une loi permettait aux préfets de concéder pour mise en culture immédiate toute parcelle de terre abandonnée ou inculte. En application de cette législation, des concessions de terres appartenant à la requérante furent décidées. Sur recours de la dame Lamotte, le Conseil d'État annula deux concessions de terre. Pour briser la résistance des juges, le législateur décida de supprimer toute possibilité de recours contre les actes de concession. Selon l'article 4 de la loi du 23 mai 1943, «l'octroi de la concession ne peut faire l'objet d'aucun recours administratif ou judiciaire». Le préfet prit un nouvel arrêté de concession et le Conseil d'État fut une nouvelle fois saisi. Le recours contentieux était-il recevable ? Il fallait, pour répondre à cette question, déterminer le sens de cette nouvelle disposition législative. Cette dernière pouvait signifier: «l'octroi de la concession ne peut faire l'objet d'aucun recours administratif ou judiciaire, y compris le recours pour excès de pouvoir qui est par conséquent irrecevable»; ou bien «l'octroi de la concession ne peut faire l'objet d'aucun recours administratif ou judiciaire, à l'exception du recours pour excès de pouvoir qui demeure recevable». Alors qu'il semblait que le législateur avait adopté la première signification, le Conseil d'État fait prévaloir la seconde: la loi «n'a pas exclu le recours pour excès de pouvoir devant le Conseil d'État contre l'acte de concession, recours qui est ouvert même sans texte contre tout acte administratif, et qui a pour effet d'assurer, conformément aux principes généraux du droit, le respect de la légalité». Quel est le véritable auteur de la norme applicable au litige, le rédacteur du texte ou son interprète ? La réponse s'impose à l'évidence: c'est l'interprète, donc le juge. Il apparaît par conséquent que, rien qu'en interprétant un texte existant, le juge dispose d'un pouvoir créateur car la norme applicable est le produit de l'interprétation du texte et non pas du texte lui-même. Or, cette interprétation est le produit d'un acte de volonté. Le juge choisit la signification qu'il souhaite donner au texte applicable au litige. Il prend pour cela une véritable décision. C'est en effet une décision «anormée» car aucune règle n'impose au juge de choisir le sens qu'il convient de donner au texte. Nous sommes alors en présence d'une véritable décision administrative. Les décisions administratives ne sont donc pas seulement celles prises par les autorités administratives. Ce sont aussi celles prises par les juges administratifs.

BIG DATA, DECISIONI AMMINISTRATIVE E “POVERTÀ” DI RISORSE DELLA PUBBLICA AMMINISTRAZIONE.

Francesco de Leonardis

SOMMARIO: 1. La prospettiva d’indagine - 2. Le tre tappe della digitalizzazione, interconnessione e intelligenza artificiale – 3. I campi di applicazione dell’analisi predittiva – 4. I vantaggi della “decisione robotica” – 5. La prima criticità: il diritto alla riservatezza - 6. La seconda criticità: l’opacità della decisione – 7. La “povertà” di risorse dell’amministrazione italiana di fronte agli algoritmi.

1. La prospettiva d’indagine.

Il tema dell’uso dei “big data” da parte delle pubbliche amministrazioni o dell’applicazione dell’intelligenza artificiale (d’ora in poi brevemente “IA”) alle amministrazioni pubbliche non appartiene più alla fantascienza, come si sarebbe potuto pensare una ventina di anni fa¹, ma è un tema di grande attualità che viene affrontato oggi dalla dottrina, e appassiona, soprattutto gli studiosi più attenti a cogliere le tendenze evolutive in atto, sia sotto l’aspetto sostanziale che processuale².

¹ Tra i primi a intuire quelli che sarebbero stati gli sviluppi tra attività amministrativa e uso della telematica non si possono non ricordare: G. DUNI, voce *Amministrazione digitale* in *Enc.dir. Annali*, vol. I, 2007 Roma, p. 13ss.; Id., *L’amministrazione digitale. Il diritto amministrativo nella evoluzione telematica*, 1992, Roma, e A. MASUCCI, *Atto amministrativo informatico* (voce), in *Enc.dir.*, Agg.to, vol. I, 2007, Milano, p. 221ss. Ovviamente senza dimenticare che lo stesso M.S. GIANNINI addirittura nel 1979 nel noto rapporto sullo stato dell’amministrazione italiana affermava che informatizzazione e sviluppo dell’amministrazione erano già allora strettamente connessi (*Rapporto sui principali problemi dell’Amministrazione dello Stato*, in *Riv. trim. dir. pubbl.*, 1982, p. 722).

² Tra i numerosi contributi in materia si possono segnalare: F. COSTANTINO, *Autonomia dell’amministrazione e innovazione digitale*, 2012, Napoli; Id., *Lampi, nuove frontiere delle decisioni amministrative*, in *Dir. amm.*, 2017, p.

Ciò che più rileva e che dimostra il rilievo non solo teorico delle questioni è il fatto che anche la giurisprudenza amministrativa ha iniziato a occuparsi della “decisione robotica”³ ossia di quella particolare decisione amministrativa adottata da un computer sulla base dell’applicazione di un algoritmo ossia di una complessa formula matematica⁴.

799 ss.; Id., *Usa della telematica nella pubblica amministrazione*, Commento all’art. 3bis della legge n. 241/90, in A. ROMANO (a cura di), *L’azione amministrativa*, 2016, Torino, p. 242 ss.; Id., *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei big data*, in *Dir.publ.*, 2019, p. 43 ss.; G. CARULLO, *Big data e pubblica amministrazione nell’era delle banche interconnesse*, in *Concorrenza e mercato*, 2016, p. 579 ss.; V. BERLINGÒ, *Datafication e giuridicizzazione*, in *Riv. trim. dir. pubbl.*, 2017, p. 641 ss.; S. CRISCI, *Intelligenza artificiale ed etica dell’algoritmo*, in *Foro amm.*, 2018, c. 1787 ss.; M. FALCONE, *Big data e pubbliche amministrazioni: nuove prospettive per la funzione conoscitiva pubblica*, in *Riv. trim. dir. pubbl.*, 2017, p. 601 ss.; Id., *Le potenzialità conoscitive dei dati amministrativi nell’era della “rivoluzione dei dati”: il caso delle politiche di eradicazione dell’epatite C*, in *Istituzioni del federalismo*, 2017, p. 27 ss.; D.U. GALETTA -J.G. CORVALAN, *Intelligenza artificiale per una pubblica amministrazione 4.0? Potenzialità, rischi e sfide della rivoluzione tecnologica in atto*, in www.federalismi.it, 3/2019; D. MARONGIU, *L’attività amministrativa automatizzata*, Rimini, 2005; A.G. OROFINO, *La patologia dell’atto amministrativo elettronico: sindacato giurisdizionale e strumenti di tutela*, in *Foro amm.-CDS*, 2002, c. 2257ss.; P. OTRANTO, *Decisione amministrativa e digitalizzazione della p.a.*, in www.federalismi.it, 2/2018; F. PATRONI GRIFFI, *La decisione robotica e il giudice amministrativo*, in www.giustiziamministrativa.it, 2018; F. SAITTA, *Le patologie dell’atto amministrativo elettronico e il sindacato del giudice amministrativo*, in *Riv. dir. amm. elettr.*, 2003, p. 1 ss.; L. VIOLA, *L’intelligenza artificiale nel procedimento e nel processo amministrativo: lo stato dell’arte*, in *Foro amm.*, 2018, c. 1598 ss.; G. AVANZINI, *Decisioni amministrative e algoritmi informatici*, Ed. Scientifica, 2019; M.L. MADDALENA, *La digitalizzazione della vita dell’amministrazione e del processo*, in *Foro amm.*, 2016.

³ È recente una decisione del Consiglio di Stato sul caso di una distribuzione di docenti effettuata da un algoritmo (Cons.Stato, VI, 8 aprile 2019, n. 2270). Ma interessanti al riguardo appaiono anche TAR Lazio, III, 21 marzo 2017, n. 3742 sulla natura vincolata dell’algoritmo che gestisce il software relativo ai trasferimenti interprovinciali del personale docente; TAR Lazio, Roma, III bis, 30 ottobre 2017, n. 10805 in tema di procedura automatizzata di classificazione delle riviste scientifiche; Cons.Stato, VI, 19 gennaio 2018, n. 353 sull’automatizzazione del criterio di concorso; TAR Lazio, II quater, 28 giugno 2016, n. 7479 in tema di contributi statali agli spettacoli dal vivo; Cons.Stato, VI, 23 gennaio 2018, n. 456 e Cons.Stato, VI, 5 dicembre 2017, n. 5733 su atti di macroorganizzazione che hanno disciplinato procedure di mobilità in forma telematica.

⁴ Sulla definizione di algoritmo come “sequenza di istruzioni computazionali che indicano ad una macchina la procedura da seguire per ottenere un certo risultato” e la definizione che di tale termine danno i manuali di in-

Il presente contributo non si propone di esaminare nel dettaglio quella che potrebbe essere definita come una delle forme più innovative di tecnificazione dell'azione amministrativa, che ha già peraltro iniziato ad essere studiata e analizzata⁵, quanto piuttosto di tentare ad inserire i big data e l'intelligenza artificiale nel quadro di evoluzione dell'amministrazione attuale per poter fornire qualche suggestione ricostruttiva.

L'analisi non può non prendere le mosse dalla nozione di "big data"⁶ che mette assieme un sostantivo "data", gli insiemi di dati elettronici, e un aggettivo "big" che fa riferimento principalmente, ma non solo, alle dimensioni quantitative di tali insiemi di dati⁷.

Si tratta di grandi insiemi di dati elettronici⁸ che lo sviluppo della tecnologia (si pensi ai computer, ma ancor più agli smartphone o ancora ai tablet etc. che fino ad una trentina di anni fa non esistevano o comunque erano prodotti di nicchia) e la sua

formatica cfr. G. AVANZINI, *Decisioni amministrative e algoritmi informatici*, cit., 6.

⁵ Sul punto si veda il contributo di V. CERULLI IRELLI, *La tecnificazione* in S. CIVITARESE MATTEUCCI–L. TORCHIA, *La tecnificazione dell'amministrazione*, in Id. (a cura di), *La tecnificazione*, in L. FERRARA–D. SORACE, *A 150 anni dall'unificazione amministrativa italiana. Studi*, vol. IV, Firenze, 2016, p. 13.

⁶ Sul punto molto interessante appare il volume *Big Data e concorrenza*, numero speciale di *Concorrenza e mercato*, a cura di F. DI PORTO, Milano, 2016 che contiene vari saggi rilevanti tra i quali F. DI PORTO, *La rivoluzione big data. Un'introduzione*, p. 5; G. CARULLO, *"Big data" e pubblica amministrazione nell'era delle banche dati interconnesse*, p. 181. Vedi anche G. COLANGELO, *"Big data", piattaforme digitali e "antitrust"*, in *Merc. conc. reg.*, 2016, p. 42; A. GIANNACCARI, *La storia dei Big Data, tra riflessioni teoriche e primi casi applicativi*, 2017, p. 307; M.F. DE TULLIO, *La "privacy" e i "big data" verso una dimensione costituzionale collettiva*, in *Pol. dir.*, 2016, p. 637; M. OREFICE, *I "Big data". regole e concorrenza*, in *Pol. dir.*, 2016, p. 697; A. MANTELERO, *Bigdata: i rischi della concentrazione del potere informativo digitale e gli strumenti di controllo*, in *Dir. informaz.*, 2012, p. 13.

⁷ Cronologicamente i big data fanno la loro comparsa a metà degli anni Novanta e a partire dal Duemila che si afferma in modo più chiaro la loro definizione che fa perno sulle quattro "V" ("Volume" che fa riferimento all'aspetto quantitativo; "Velocity" che fa riferimento alla rapidità con la quale i dati posso essere generati e trasmessi; "Variety" che fa riferimento alla diversità dei dati e della loro origine; "Veracity" che fa riferimento alla qualità, correttezza e affidabilità dei dati): sul punto v. C. COMELLA, *Origine dei "Big data"*, in *Gnosis*, 2017.

⁸ Ognuno di noi genera quotidianamente centinaia di dati, se si mettono insieme i milioni di persone che abitano il nostro paese si generano miliardi di dati e se si mettono insieme i dati prodotti da cittadini dello stesso continente i numeri diventano talmente grandi da non potere essere menzionati.

diffusione a costi accettabili per quasi tutti (oggi è difficile trovare chi non abbia un telefono mobile o non usi la rete) ha reso disponibili sia ad operatori privati che ad operatori pubblici⁹.

In questo contesto la pubblica amministrazione si pone come un punto di riferimento chiave dal momento che essa costituisce in assoluto il maggior detentore di dati (si pensi alle varie anagrafi civile, sanitaria, fiscale, al catasto, alla motorizzazione civile, alle banche dati del Ministero degli Interni sulla violazioni del codice della strada o per le indagini penali etc.)¹⁰ ed è dunque comprensibile che il legislatore si sia occupato fin da subito della regolazione dei rapporti tra pubblica amministrazione e big data.

2. Le tre tappe della digitalizzazione, interconnessione e intelligenza artificiale.

I rapporti tra big data e amministrazione possono essere sintetizzati in tre tappe logiche fondamentali: quella che si potrebbe definire la tappa della “digitalizzazione”, quella che si potrebbe definire come l’“interconnessione dei dati” e, infine, quella in cui ci troviamo che è quella dell’applicazione dell’IA alle decisioni amministrative ossia, come viene spesso menzionata, la tappa dell’“analisi predittiva”.

Posto che come si è detto in apertura le amministrazioni sono storicamente detentori di una pluralità di dati di vario genere rispetto ai cittadini, alle imprese, all’ambiente, all’economia etc. e che tali dati venivano detenuti (e spesso continuano ad esserlo) in modo cartaceo, il primo passaggio logico affrontato dal legislatore è stato quello della dematerializzazione dei dati¹¹ ossia, in parole semplici, di trasformare i documenti cartacei in *file*.

La spinta forte alle amministrazioni nel senso di raccogliere,

⁹ Chiunque, consapevolmente o inconsapevolmente, lascia una traccia digitale delle mille e mille azioni che compie ogni giorno nella vita quotidiana: si pensi non solo alle mail che inviamo, ai messaggi telefonici, a whatsapp ma anche ai siti web che visitiamo, alle telecamere che c’inquadrano mentre facciamo la spesa, alle telecamere che ci seguono per le strade etc.

¹⁰ Lo ricorda il considerando 4 della direttiva 2003/98.

¹¹ Si noti che già il d.lgs. 3 febbraio 1993, n. 29, nell’occuparsi della riforma del pubblico impiego, affrontava, seppur nell’ottica dell’efficientamento delle strutture organizzative, il nodo dell’informatizzazione degli uffici pubblici: F. COSTANTINO, *Informatizzazione della p.a.*, cit., 2015.

mappare o rilevare i dati direttamente in formato digitale è contenuta nel codice dell'amministrazione digitale (CAD), il d.lgs. 7 marzo 2005, n. 82¹², che, anche sulla spinta del diritto europeo, ha regolato ad ampio spettro i possibili impieghi pubblici dell'informatizzazione, raccogliendo la precedente normativa anche in materia di documentazione.

Tale digitalizzazione ha comportato una serie di vantaggi tra i quali i seguenti: un primo vantaggio di tipo logistico è quello che occorre molto meno spazio per conservare la documentazione; un secondo vantaggio attiene alla sicurezza e alla conservazione (i file possono essere conservati in luoghi diversi e posti al riparo ad esempio da catastrofi naturali etc.)¹³; un terzo vantaggio è quello della fruibilità (all'amministrazione viene chiesto di porre a disposizione della collettività l'enorme patrimonio conoscitivo in suo possesso consentendo ai cittadini e alle imprese di sfruttare i dati per molteplici scopi)¹⁴; un quarto vantaggio è quello del "ridare fiducia" all'amministrazione (proprio tale trasparenza si immagina possa accrescere la fiducia dei cittadini nell'amministrazione in quanto permette un loro maggiore coinvolgimento nel cambiamento organizzativo e, conseguentemente, finisce con il responsabilizzare l'amministrazione).

Sotto tale ultimo aspetto non si può però non ricordare che la digitalizzazione e la connessa possibilità di rendere pubblici

¹² Volendo citare le norme essenziali in materia non si possono non menzionare anche le cd. Riforme Brunetta del 2009-2010 che hanno collegato efficientamento dell'amministrazione e informatizzazione; il d.lgs. 14.3.2013, n. 33 sulla Trasparenza e le misure per l'attuazione della cd. Agenda digitale.

¹³ Si tratta delle soluzioni di *cloud computing* attraverso le quali i dati vengono archiviati o elaborati da applicazioni e terminali virtuali, che non sono fisicamente presso l'amministrazione, ma su server che possono fisicamente anche trovarsi in luoghi enormemente distanti: sul punto v. *Raccomandazioni e proposte sull'utilizzo del cloud computing nella pubblica amministrazione*, DigitPA 8 giugno 2012.

¹⁴ Si tratta del cd. riutilizzo in relazione all'uso per fine diverso da quello per il quale il dato è stato prodotto (art. 2, lett. e), d.lgs. 24.1.2006, n. 36, e dei cd. open data (dati in formato aperto). Il sito italiano di riferimento per gli open data è *dati.gov.it*. L'attenzione per gli open data è di livello anche internazionale: si veda la direttiva del Presidente Obama del dicembre 2009, seguita dall'istituzione del sito *data.gov*, e dall'executive order del 9 maggio 2013; si veda l'organizzazione Open Government Partnership; la Carta adottata dal G8; l'italiano Open data action plan. A livello europeo del riutilizzo l'Europa si era occupata già con la direttiva 2003/98/CE modificata dalla dir. 2013/37/UE. Sul punto cfr. F. COSTANTINO, *Informatizzazione della p.a.*, cit., 2015.

in rete una serie di dati e di documenti non è di per sé sinonimo di trasparenza perché il risultato di non far passare la conoscenza dall'amministrazione ai cittadini si può ottenere paradossalmente, oltre che attraverso il segreto, anche inviando al destinatario una tale valanga di informazioni che gli renda difficile comprendere il cuore del problema¹⁵.

Il secondo passaggio logico è quello dell'interconnessione dei dati o del sistema pubblico di connettività¹⁶: solo riuscendo a combinare, a mettere insieme la mole enorme di dati detenuti dalla p.a., appunto attraverso l'applicazione di formule matematiche, si riesce ad avere un insieme di dati gestibile come sistema e non vari insiemi monadici di dati che non parlano tra di loro.

La stessa strategia Europa 2020¹⁷ in vigore dal 2010, tra le varie azioni, aree e obiettivi richiede di istituire un mercato digitale unico, di favorire l'interoperabilità (ossia lo scambio e l'interazione dei dati) e l'uso di standard, di migliorare la fiducia in internet e la sicurezza online, di aumentare la velocità di accesso alla rete, di potenziare gli investimenti per la ricerca e lo sviluppo, di migliorare l'alfabetizzazione e l'inclusione digitale, di applicare le ICT.

I due precedenti passaggi sono strumentali all'obiettivo finale: quello dell'applicazione dell'IA ai big data che sono stati interconnessi; è qui che entra in gioco l'algoritmo¹⁸ che in que-

¹⁵ Così rispetto al supporto cartaceo, è possibile senza maggiori costi la pubblicazione di quantità enormi di informazioni, e troppe informazioni, o informazioni che prese da sole siano incomprensibili, o artatamente selezionate od orientate in una specifica destinazione, hanno l'effetto di occultare più che rivelare: cfr. E. CARLONI, *L'amministrazione aperta: regole strumenti limiti dell'open government*, Rimini, 2014, p. 29.

¹⁶ Su tali aspetti si veda: Comunicazione della Commissione 6 maggio 2015, *Strategia per il mercato unico digitale in Europa*, COM (2015) 192, finale. Il tema della connessione è presente ovviamente nel Codice dell'amministrazione digitale, si v. ad es. l'art. 60 e l'art. 58; si v. anche F. CARDARELLI, *Uso della telematica, commento all'art. 3 bis della legge 241/90* in M.A. SANDULLI (a cura di), *Codice dell'azione amministrativa*, Milano, 2017, p. 427 e F. COSTANTINO, *L'uso della telematica nella Pubblica Amministrazione*, in *L'azione amministrativa*, a cura di Alberto Romano, 2016, Milano.

¹⁷ COM/2010/0245 f/2.

¹⁸ Come è stato intelligentemente rilevato nel diritto positivo viene utilizzato il termine algoritmo solo in discipline speciali mentre ampio è il riferimento ad altre terminologie come quella di amministrazione digitale, di diritti di cittadinanza digitale, di documento informatico o di decisione automatizzata: G. AVANZINI, *Decisioni amministrative e algoritmi informatici*, cit., 2.

sto ambito funziona come una sorte di chiave che apre alla p.a. la possibilità di prevedere i comportamenti futuri.

È proprio l'IA o l'algoritmo che consente di "far parlare" i dati attraverso le ricerca di correlazioni¹⁹: si tratta dei cd. algoritmi predittivi e di raccomandazione che consentono di effettuare previsioni di comportamenti futuri partendo dall'analisi di comportamenti passati ben noti alle imprese e operatori privati che si occupano di marketing.

Il nostro ordinamento contiene due disposizioni fondamentali al riguardo: l'art. 50 CAD dispone che le pubbliche amministrazioni, nell'ambito delle proprie funzioni istituzionali, procedano all'analisi dei propri dati, anche in combinazione con altre amministrazioni (o gestori di servizi pubblici per profili di pubblico interesse o società a controllo pubblico) e l'art. 50-ter (introdotto dal d.lgs. 13 dicembre 2017, n. 217) che istituzionalizza un progetto di piattaforma digitale nazionale dei dati (PDND)²⁰.

Molto importante è anche il Libro Bianco adottato per l'Italia Digitale (AGID) dedicato all'intelligenza artificiale pubblicato nel 2018 e che, significativamente, considera tale istituto quale strumento di sviluppo sociale, economico e culturale.

Com'è stato correttamente rilevato "l'uso dei big data in chiave previsionale consente alle amministrazioni, anche locali, di adottare un diverso modo di operare. Non solo esse possono, infatti, investire e impiegare le risorse – sempre più limitate – principalmente ove si produrrà il maggiore impatto favorevole. Più a monte, sono in grado di riconfigurare i servizi, prevedendo problemi, anticipando bisogni, in ultima analisi pianificando in maniera strategica. In tal modo esse sono in grado di evitare di dovere affrontare rischi ed intervenire principalmente in occasione di inconvenienti, con conseguenze rilevanti in termini di costi. Si tratta perciò di un mutamento epo-

¹⁹ G. AVANZINI, *Decisioni amministrative e algoritmi informatici*, cit., 11.

²⁰ Anche il Piano triennale per l'informatica nella Pubblica Amministrazione 2017-2019 si occupa di intelligenza artificiale anche se rinvia la sua attuazione proprio alla realizzazione della struttura ritenuta indispensabile per lo sfruttamento dei big data, la Piattaforma Digitale Nazionale Dati (PDND) a cui viene affidato il compito di valorizzare il patrimonio informativo della PA mediante l'utilizzo delle tecnologie big data al servizio dei decisori pubblici, grazie all'analisi esplorativa dei dati.

cale per istituzioni che generalmente agiscono solo per rispondere agli stimoli”²¹.

È dunque proprio attraverso l’interconnessione e la lavorazione del dato o meglio dell’insieme dei dati che si perviene alla conoscenza dei fenomeni e dei loro andamenti e quindi al cd. *government by data*.

E dunque non solo l’amministrazione digitale, grazie alle ICT, può disporre di una base conoscitiva di dati fino a pochi anni or sono inimmaginabile, attraverso banche dati condivise, interoperabili e consultabili con modalità telematica ma tali dati possono essere messi in relazione, appunto, da formule matematiche al fine di produrre delle vere e proprie decisioni amministrative.

3. I campi di applicazione dell’analisi predittiva.

I concetti di cui si sta parlando sono realtà già da tempo in una pluralità di settori in cui i dati vengono rilevati, interconnessi e lavorati dalle pubbliche amministrazioni per finalità di interesse generale. Può essere utile passarne in rassegna i principali.

Si pensi innanzitutto alla tutela in materia di sicurezza esterna: sono diffuse e comuni le cd. “no fly list” che consentono di prevedere e pianificare i controlli di sicurezza e per le quali si pongono non pochi problemi (ad es. la discriminazione operata in base all’appartenenza ad un determinato gruppo etnico).

Si pensi ancora alla tutela in materia di sicurezza interna: in una cittadina degli Stati Uniti (Memphis) analizzando la frequenza, i luoghi e gli orari dei reati, i crimini si sono ridotti di circa il 25% perché l’algoritmo utilizzato consentiva alle forze dell’ordine di prevedere i luoghi dove tale crimini sarebbero stati realizzati²² e allo stesso modo a Los Angeles analizzando i dati rilevati in pochi mesi i reati si sono ridotti del 15%²³.

²¹ F. COSTANTINO, *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei big data*, cit., 2019, p. 48.

²² R. IBM, *Analytics Help Memphis Cops Get “Smart”*, <http://www.bloomberg.com/news/articles/2011-12-05/ibm-analytics-help-memphis-cops-get-smart>.

²³ C. KAHN, AT LAPD, *Predicting Crimes Before They Happen*, <http://www.npr.org>. Cfr. anche A. G. FERGUSON, *Big Data and Predictive Reasonable Suspicion*, in *U. Pa. L. Rev.*, 2015, 410.

Vi è ancora la materia sanitaria: analizzando le ricerche fatte da utenti di Google sui sintomi di certe malattie si è riusciti a prevedere la diffusione di un'epidemia (si tratta del progetto Flu trends)²⁴; si possono gestire meglio le liste di attesa degli ospedali²⁵; si possono effettuare campagne di prevenzione; si possono individuare gli effetti collaterali negativi rispetto a determinati farmaci²⁶.

Un campo di elezione per l'uso dei big data da parte della p.a. è quello della programmazione economica: le previsioni economiche che determinano politiche governative sono più accurate se si basano su dati certi e la complessità delle operazioni di previsione viene di molto facilitata dall'uso delle macchine.

Vi sono i contratti pubblici con le aste elettroniche e le gare telematiche, le sanzioni automatizzate del codice della strada, la SCIA automatizzata, il rating di impresa fornito dall'ANAC, la mobilità del personale dipendente e le procedure preselettive²⁷.

Ma vi è anche il settore ben noto a chi si occupa di università della programmazione universitaria: si pensi alle valutazioni VQR o VTR che vengono utilizzate dagli atenei per mettere a concorso o meno un posto; la ripartizione del FFO tra le università; le mediane che consentono di diventare commissario ASN o di partecipare ad una procedura di abilitazione... in tutti i casi la presenza dell'algoritmo è silenziosa e ingombrante.

Vi è poi la materia dell'assetto del territorio: le analisi del

²⁴ Sulla base delle ricerche fatte dagli utenti sui sintomi influenzali è stato possibile prevedere dove si sarebbe diffusa l'epidemia, in quanto i malati consultavano Google prima di recarsi presso un presidio medico. Sul punto: M. MATTIOLI, *Disclosing Big Data*, in *Minn. L. Rev.*, 2014, p. 535, in part. 540.

²⁵ L'impatto dei big data in materia sanitaria è stato analizzato da B.R. FURROW, *Searching for Adverse Events: Big Data and Beyond*, in *Annals Health L.*, 2018, p. 149. Si è osservato che attraverso l'analisi predittiva si potenzia la capacità di identificare i pazienti ad alto rischio (con riduzione dei costi del reparto di emergenza), si possono ridurre il numero di visite e offrire assistenza personalizzata. Inoltre, gli ospedali possono analizzare i tassi di ammissione e assegnare il personale adeguato in base alle previsioni, con una riduzione dei costi ospedalieri e dei tempi di attesa.

²⁶ Dalle ricerche effettuate su internet relative a due farmaci in relazione all'ipoglicemia gli studiosi hanno ipotizzato che potesse esserci una correlazione tra l'uso dei due farmaci e un effetto collaterale, e le prove sperimentali hanno poi confermato l'ipotesi. Cfr. C. ANDERSON, *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, *Wired.com/ 2008/06/pb-theory/*.

²⁷ Su questi temi si rinvia al completo lavoro monografico di G. AVANZINI, *Decisioni amministrative e algoritmi informatici*, 2019, cit., p. 41 ss.

territorio che consentono la programmazione urbanistica o la gestione del traffico (attraverso telecamere, sensori si può gestire il traffico, l'illuminazione e la gestione dei servizi)²⁸.

Si pensi alla gestione dei servizi: prevedere il carico della rete elettrica costituisce un esempio di analisi predittiva da cui discendono decisioni in ordine a come usare le centrali elettriche²⁹.

Si pensi ancora alla previdenza sociale: il caso INPS in cui il sistema di big data è stato realizzato fin dal 2000 a supporto delle decisioni strategiche dell'Istituto. Il sistema acquisisce anche le informazioni degli ispettori di vigilanza utile in particolare per l'attività di "intelligence antifrode", che rileva i casi di doppia erogazione degli assegni familiari all'interno della stessa famiglia, di indebito conguaglio dell'indennità di malattia ecc.

Rilevantissimo campo dell'uso dei big data è quello della pubblica incolumità: si pensi solo alle previsioni metereologiche sulla base delle quali vengono adottate le ordinanze comunali o in base alle quali si dirama l'allerta di vari gradi in varie zone.

Vi è poi il tema dei controlli che come si dirà nel paragrafo finale potrebbe costituire uno dei temi di maggiore sviluppo dei big data nella p.a.

E, infine, *last but not least* la materia ambientale dove il diritto di accesso trova la sua massima esplicazione: nel 2015 il premio Best Policy Insights Hack è stato attribuito a un software privato che indicava, grazie ai big data, in quali quartieri fosse preferibile installare pannelli solari³⁰.

Si tratta in sintesi per le p.a. di immagazzinare dati e di lavorarli: se in molti casi tali lavorazioni consentono di perseguire l'interesse pubblico in modo migliore di prima, d'altra parte non

²⁸ Con riferimento alle smart cities si v. Y. LV, Y. DUAN, W. KANG et al., *Traffic Flow Prediction With Big Data: A Deep Learning Approach*, *Ieee transactions on intelligent transportation systems*, 2015, p. 865; il n. 5 di *Istituzioni del federalismo*, 2015; R. FERRARA, *The smart city and the green economy in europe: a critical approach*, in *Dir. proc. amm.*, 2015, fasc. 2-3, p. 63; E. FERRERO, *Le "smart cities" nell'ordinamento giuridico*, in *Foro amm.*, 2015, p. 1267; A. CASINELLI, *Le città e le comunità intelligenti (Commento a d.l. 18 ottobre 2012, n. 179)*, in *Giorn. dir. amm.*, 2013, p. 240.

²⁹ M. PATERSON, M. McDONAGH, *Data Protection in an Era of Big Data: The Challenges Posed by Big Personal Data*, 2018, in *Monash U. L. Rev.*, 1, p. 6.

³⁰ F. COSTANTINO, *Intelligenza artificiale e decisioni amministrative*, in *Riv.it.sc.giur.*, 2017, p. 370.

sono pochi i rischi e gli svantaggi che l'analisi predittiva comporta.

4. I vantaggi della “decisione robotica”.

La decisione robotica, quella presa da una macchina/elaboratore/computer, almeno a prima vista, ha diversi pregi: è poco costosa da un punto di vista economico³¹; è molto rapida e risponde ad esigenze di semplificazione³²; riduce se non annulla la discrezionalità della p.a.; appare imparziale (una macchina non può favorire l'uno piuttosto che l'altro); consente un ampliamento della partecipazione in altri tempi e contesti inimmaginabile³³; è affidabile e certa³⁴.

³¹ È noto che una macchina costa meno del lavoro dell'uomo e che i costi maggiori di un'impresa ma anche della pubblica amministrazione stessa sono quelli del personale. È stato infatti giustamente messo in evidenza che la capacità di risparmio è il criterio privilegiato da tenere in conto per la decisione circa gli investimenti in informatizzazione (art. 15, co. 2-ter, d.lgs. 8.3.2005, n. 52, Codice dell'Amministrazione Digitale - CAD, introdotto dal d.lgs. 30.12.2010, n. 235): F. COSTANTINO, *Informatizzazione della P.a.*, in www.treccani.it, 2015.

³² Le cd. leggi Bassanini (l. 15 marzo 1997, n. 59 ss.) che hanno in particolare disciplinato la comunicazione tra p.a. e cittadini hanno interpretato espressamente l'informatizzazione come strumento di semplificazione. Sul rapporto tra semplificazione ed interazione telematica tra cittadino ed amministrazione, cfr. P. LAZZARA, *Principio di semplificazione e situazioni giuridiche soggettive*, in *Dir. amm.*, 2011, p. 679 ss. Sulla rapidità della decisione è significativo l'esempio riportato dal bel saggio di D.U. GALETTA per cui il lavoro richiesto per avviare una gara di appalto di beni e servizi completamente digitalizzati nella città di Buenos Aires, come l'acquisto di “accessori per computer”, prima implicava una media di 29 giorni lavorativi, di 670 clic, l'apertura di circa 60 finestre ed il “copia e incolla” di numerosi dati ed oggi solo quattro minuti: D.U. GALETTA - J.G. CORVALAN, *Intelligenza artificiale per una pubblica amministrazione 4.0? Potenzialità, rischi e sfide della rivoluzione tecnologica in atto*, in www.federalismi.it, 3/2019, cit.

³³ Si pensi a quanto le tecnologie informatiche rendano più praticabili vere e proprie forme di democrazia diretta (lo stesso art. 9 CAD prevede che l'ICT favorisca l'esercizio dei diritti politici e civili individuali e collettivi): cfr. sul punto S. RODOTÀ, *Iperdemocrazia*, 2013, Roma-Bari; M.A. SHAREEF - N. ARCHER, *Mission of Transformational Government as a Postmodern Organisation: A critical theory perspective*, in *Transformational government Through egov practice: socioeconomic, cultural, and technological issues*, 2012, p. 516.

³⁴ Sul punto però si veda il comprensibile *caveat* di F. COSTANTINO, *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei big data*,

Si può ricordare con riferimento a quest'ultimo punto che, ad esempio, per fare l'analisi di rischio di un terreno per verificare se sia inquinato o meno basta inserire i dati in un software che consente all'impresa di sapere direttamente se il suo lavoro di bonifica può dirsi terminato oppure no³⁵.

La decisione assunta da un IA e fondata sui big data è, quindi, assai più prevedibile di quella assunta da un funzionario della p.a. e ciò risponde al forte desiderio delle imprese di prevedibilità e di certezza: i risultati cui perviene l'analisi algoritmica tendono ad essere prevedibili e precisi così rispondendo al principio di certezza del diritto³⁶.

Se ciò s'inquadra nell'attuale trend di riduzione della discrezionalità delle p.a. di cui si parlerà nella parte finale non si dubita che ciò possa senz'altro costituire un elemento a favore dell'uso dell'IA nell'ambito dell'attività amministrativa.

Nella stessa linea potrebbe giocare il clima generale di sospetto e di sfiducia nella pubblica amministrazione: dato che il funzionario pubblico nell'immaginario collettivo è tendenzialmente soggetto a pressioni esterne, potrebbe sostenersi facilmente che la sua sostituzione con una macchina potrebbe dimi-

cit., 2019, p. 55, per il quale “non si deve neppure pensare che gli algoritmi che scandagliano intrecciano ed elaborano i big data siano neutrali. La mano del programmatore incide su come opera l'algoritmo, e ognuno di essi è solo uno dei possibili: in questo senso diventa importante il rapporto tra scienziato e decisore pubblico. La supposta neutralità tecnologica è infatti un classico espediente per eludere il problema dell'accountability democratica del decisore. Anche per queste ragioni, alcuni “pregiudizi” sottostanti agli algoritmi non possono mai del tutto essere rimossi, anche perché non è detto che siano esplicitati, e non necessariamente per dolo, ma generalmente per inconsapevolezza: come effetto, vengono replicati i valori dei cd. scienziati dei dati, con il rischio di perpetuare visioni della società frutto di stereotipi e pregiudizi”.

³⁵ Si tratta dei software Risk-net, vers. 2.1 e vers. 3.1 sviluppati da RECONnet - Rete Nazionale sulla gestione e la Bonifica dei Siti Contaminati con l'obiettivo di fornire uno strumento che ricalchi la procedura APAT-ISPRA di Analisi di Rischio (“Criteri metodologici l'applicazione dell'analisi assoluta di rischio ai siti contaminati”; APAT 2008) in accordo con quanto previsto dalla normativa italiana (D. Lgs. 152/06 e D. Lgs. 04/08).

³⁶ Tra gli studiosi di diritto amministrativo il tema della prevedibilità delle scelte è stato affrontato in numerosissimi saggi tra di essi si può segnalare: A. POLICE, *Prevedibilità delle scelte e certezza dell'azione amministrativa*, in *Dir.amm.*, 1996, p. 697; M. IMMORDINO, *Certezza del diritto e amministrazione di risultato*, in *Principio di legalità e amministrazione di risultati*, Torino, 2004. Evidenzia, ad esempio, che la produzione di certezze costituisce una delle funzioni fondanti dei pubblici poteri: A. FIORITTO, *La funzione di certezza pubblica*, Padova, 2003, p. 396.

nuire, se non addirittura in alcuni casi azzerare, il fenomeno della corruzione.

5. La prima criticità: il diritto alla riservatezza.

La lavorazione dei big data da parte dell'IA al fine di pervenire ad una decisione robotica si deve confrontare, però, con un primo grande problema che è quello della riservatezza³⁷.

Ed infatti perché la previsione sia il più possibile precisa occorre disporre e lavorare un'enorme quantità di dati: se non si richiede il consenso degli interessati (come ad esempio avviene a Singapore) ovviamente la quantità di dati a disposizione delle amministrazioni sarà massima e maggiore sarà la capacità predittiva; se invece si richiede il consenso l'insieme dei dati consentirà più difficilmente di svolgere l'analisi predittiva.

Si tratta di trovare il bilanciamento giusto e al riguardo il diritto europeo si è pronunciato con il Regolamento 2016/679/

³⁷ In generale sul tema della riservatezza S. CALZOLAIO, *Protezione dei dati personali (voce)*, in *Disc pubbl.*, Agg.to, 2017, Torino, p. 594 ss.; ID., *Privacy by design. Principi dinamiche ambizioni del nuovo Reg. UE 2016/679*, in *www.federalismi.it*. Sui rapporti tra trasparenza e riservatezza cfr. B. CAROTTI, *L'amministrazione digitale e la trasparenza amministrativa*, in *Giorn.dir.amm.*, 2015, p. 625; M.C. CAVALLARO, *Garanzie della trasparenza amministrativa e tutela dei privati*, in *Dir.amm.*, 2015, p. 121; M.C. D'Arienzo, *Diritto alla trasparenza e tutela dei dati personali nel d.gs. n. 33/2013, con particolare riferimento alla disciplina dell'accesso civico*, in *Dir.processo amm.*, 2015, p. 123; F. FRACCHIA, *L'impatto delle misure anticorruzione e della trasparenza sull'organizzazione amministrativa*, in *Il dir.dell'economia*, 2015, p. 483; G. VESPERINI (a cura di), *Il big bang della trasparenza*, 2015; A. PAJNO, *Il principio di trasparenza alla luce delle norme anticorruzione*, in *Giust.civ.*, 2015, p. 213; M.R. SPASIANO, *Riflessioni in tema di trasparenza anche alla luce del diritto di accesso civico*, in *Nuove autonomie*, 2015, p. 63; S. VACCARI, *Il difficile bilanciamento tra favor per la trasparenza e (necessaria) tutela della riservatezza nel d.lgs. n. 33/2013*, in *Il dir.dell'economia*, 2015, p. 151; A. CONTIERI, *Trasparenza e accesso civico*, in *Nuove autonomie*, 2014, p. 563; Atti del convegno di Varenna, *Politica e amministrazione della spesa pubblica: controlli, trasparenza e lotta alla corruzione*, 2014 F.II.593 59; F. MANGANARO, *Trasparenza e obblighi di pubblicazione*, in *Nuove autonomie*, 2014, p. 553; I. NICOTRA, *La dimensione della trasparenza tra diritto alla accessibilità totale e protezione dei dati personali: alla ricerca di un equilibrio costituzionale*, in *Studi in onore di Maurizio Pedrazza Gorlero*, vol. II, *La libertà d'informazione e la democrazia costituzionale*, 2014, p. 429 ss.; F. PATRONI GRIFFI, *La trasparenza della pubblica amministrazione tra accessibilità totale e riservatezza*, in *Federalismi*, n. 8, 2013.

UE, cd. GDPR (General Data Protection Regulation), entrato in vigore il 25 maggio 2018, con cui si potrebbe dire si dà un colpo al cerchio e uno alla botte ai due interessi coinvolti del buon andamento e della tutela della riservatezza³⁸.

Si prevede che l'accesso ai dati in possesso di altre amministrazioni debba avvenire sempre sulla base di richieste scritte, motivate, occasionali, non possa riguardare un intero archivio o condurre all'interconnessione di archivi.

Il Regolamento UE 2016/679 o GDPR relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati) e, di conseguenza, sostituisce il Codice Privacy italiano (D.Lgs. 196 del 2003)³⁹.

Il principio espresso dalla normativa europea per le p.a., in particolare nei considerando 31, 111, 154, 47 non è quello della interconnessione sempre e dovunque tra le p.a. ma quello della interconnessione prevista solo per specifici casi o per specifiche funzioni.

Nel considerando 71 è enunciato il principio secondo il quale l'interessato "dovrebbe avere il diritto di non essere sottoposto a una decisione, che possa includere una misura, che valuti aspetti personali che lo riguardano, basata unicamente su un trattamento automatizzato e che produca effetti giuridici che lo riguardano o incida in modo analogo significativamente sulla sua persona".

In particolare ci si sofferma sulla "profilazione", ossia sul trattamento automatizzato che valuta aspetti personali, "al fine di analizzare o prevedere aspetti riguardanti il rendimento pro-

³⁸ Dei rapporti tra big data e decisione amministrativa si occupa un'indagine conoscitiva avviata da parte dell'Autorità Garante della Concorrenza e del Mercato (AGCM), dell'Autorità per le Garanzie e nelle Comunicazioni (AGCOM) e dell'Autorità Garante per la protezione dei dati personali (Garante della Privacy), che si è concentrata in particolare sui profili della concorrenza e della riservatezza.

³⁹ Il decreto 196 del 2003 "Codice in materia di protezione dei dati personali" detto anche "Testo unico sulla Privacy" è la normativa entrata in vigore dal 1° gennaio 2004, ampliando il percorso legislativo compiuto dall'Italia in materia di dati personali a partire dalla legge 675/96. Con tale Codice si è stabilito che ciascun soggetto (professionista, azienda, ente, associazione, ecc.) che tratti i dati personali di persone fisiche terze debba adottare misure tecniche e obblighi (misure minime) a garanzia di standard di sicurezza minimi.

fessionale, la situazione economica, la salute, le preferenze o gli interessi personali, l'affidabilità o il comportamento, l'ubicazione o gli spostamenti dell'interessato, ove ciò produca effetti giuridici che la riguardano o incida in modo analogo significativamente sulla sua persona”.

A seconda delle circostanze e del contesto, ci si raccomanda affinché siano adottate procedure matematiche o statistiche appropriate, e siano garantite la rettifica delle inesattezze dei dati, la minimizzazione del rischio di errori, la sicurezza dei dati e l'insussistenza di fenomeni di discriminazione.

Nell'art. 5 del regolamento GDPR si prevede, nel rispetto del principio della minimizzazione dei dati, che i dati personali siano raccolti e trattati per finalità determinate, esplicite e legittime, e che siano adeguati, pertinenti e limitati a quanto necessario rispetto alle finalità.

Nell'art. 6 dello stesso si afferma però che sia lecito il trattamento dei dati necessario per adempiere un obbligo legale o per l'esecuzione di un compito di interesse pubblico o connesso all'esercizio di pubblici poteri: si rende così possibile il trattamento per una finalità diversa da quella per la quale i dati personali sono stati raccolti senza consenso dell'interessato e senza atto legislativo, ma il trattamento deve essere compatibile con la finalità originaria (tenendo conto dei nessi tra le finalità, del contesto, della natura dei dati, delle possibili conseguenze, delle garanzie offerte).

Si noti che, in base all'art. 9, è possibile trattare anche dati c.d. sensibili o super sensibili per motivi di interesse pubblico rilevante, purché il trattamento sia proporzionato alla finalità, rispetti l'essenza del diritto alla protezione dei dati e preveda misure appropriate e specifiche per tutelare i diritti fondamentali e gli interessi dell'interessato, ma anche se il trattamento sia necessario per finalità di medicina preventiva o di medicina del lavoro, valutazione della capacità lavorativa del dipendente, diagnosi, assistenza o terapia sanitaria o sociale ovvero gestione dei sistemi e servizi sanitari o sociali, nel caso di protezione da gravi minacce per la salute a carattere transfrontaliero o la garanzia di parametri elevati di qualità e sicurezza dell'assistenza sanitaria e dei medicinali e dei dispositivi medici.

Se i dati non siano stati ottenuti presso l'interessato, l'informazione allo stesso, anche relativa all'esistenza di un processo decisionale automatizzato, compresa la profilazione e informazioni significative sulla logica utilizzata, nonché l'importanza

e le conseguenze previste di tale trattamento non si applicano se e nella misura in cui comunicare tali informazioni risulti impossibile o implicherebbe uno sforzo sproporzionato, qualora l'ottenimento o la comunicazione siano espressamente previsti dal diritto dell'Unione o dello Stato membro che prevede misure appropriate per tutelare gli interessi legittimi dell'interessato; oppure qualora i dati personali debbano rimanere riservati conformemente a un obbligo di segreto professionale.

Come è stato esattamente rilevato⁴⁰ i primi atti successivi al Regolamento, quali la Direttiva 2016/680, relativa al trattamento dei dati personali a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, e la Direttiva 2016/681, sull'uso dei dati del codice di prenotazione (PNR, Passenger Name Record) a fini di prevenzione, accertamento, indagine e azione penale nei confronti dei reati di terrorismo e dei reati gravi hanno messo a dura prova la nuova disciplina rispetto all'esigenza di lotta al crimine e alla normazione di altri ordinamenti.

Al riguardo il Garante europeo per la protezione dei dati personali ha significativamente osservato in data 14 marzo 2017 che "la direttiva trasforma tutti noi in sospettati. I dati verranno analizzati in base ad algoritmi di cui non conosciamo il funzionamento. Tutti saranno schedati, anche coloro che non hanno nulla a che fare con la prevenzione di reati. L'Europa getta così una rete in termini troppo larghi. Abbiamo dubbi sul fatto che la proporzionalità sia stata rispettata al cento per cento".

6. La seconda criticità: l'opacità della decisione.

La decisione amministrativa robotica ha due grandi difetti: non consente a chi debba sindacarne la legittimità di ricostruirne l'iter logico giuridico e viola il principio di trasparenza⁴¹.

Tali due difetti vengono comunemente sintetizzati nel concetto di opacità della decisione che è aggravata dal fatto che molti prodotti predittivi vengono sviluppati da aziende private

⁴⁰ F. COSTANTINO, *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei big data*, cit., 2019.

⁴¹ Sul tema v. F. PATRONI GRIFFI, *La decisione robotica e il giudice amministrativo*, in <https://www.giustizia-amministrativa.it> (28.08.2018), p. 5.

dal momento che le amministrazioni, sempre più spesso, non hanno competenze e mezzi per generare e gestire modelli di previsione per l'allocazione di servizi pubblici rilevanti.

Come è stato esattamente rilevato il rischio, in queste ipotesi, è che i soggetti pubblici rimangano all'oscuro dello sviluppo e dell'implementazione di questi modelli e che addirittura alle richieste delle amministrazioni si possa opporre l'eccezione di segretezza industriale⁴².

Si potrebbe dire che la decisione amministrativa robotica di per sé violi innanzitutto il dovere di motivazione del provvedimento amministrativo dal momento che in essa non si evincono né le ragioni giuridiche né i presupposti di fatto di tale decisione: così ad esempio viene riportato l'esempio del licenziamento di una serie di professori negli USA sulla base dei test sostenuti dagli studenti in cui l'algoritmo non consentiva agli insegnanti di comprendere le ragioni di tale provvedimento.

Allo stesso modo la decisione robotica viola anche il principio di trasparenza e quello di pubblicità che sono diventati uno delle linee di evoluzione trainante della legislazione dell'ultimo ventennio.

In un caso si era posto il problema se fosse legittimo accedere all'algoritmo del programma impiegato dal M.I.U.R. per lo svolgimento del procedimento relativo al trasferimento interprovinciale del personale docente.

Il giudice di primo grado aveva affermato la piena legittimità della scelta (discrezionale) dell'amministrazione di gestire il procedimento attraverso un "atto ad elaborazione elettronica" in cui "l'elaborazione del contenuto dell'atto viene affidata interamente allo strumento informatico e, quindi, in definitiva alla macchina, la quale provvede direttamente al reperimento, al collegamento e alla interrelazione tra norme e dati assumendo, conseguentemente, un ruolo strumentale rispetto all'atto amministrativo conclusivo. Nella predetta fattispecie è l'elaborazione stessa del contenuto dell'atto che si svolge elettronicamente, elaborazione che consiste, appunto, nello svolgimento dell'iter logico che conduce alla redazione dell'atto finale in relazione al rispettivo contenuto e che concretizza la sua motivazione. Il documento finale che contiene la predetta elaborazione, invece,

⁴² F. COSTANTINO, *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei big data*, cit., 2019, p. 53.

può avere qualsiasi forma ammessa dall'ordinamento e, quindi, essere anche cartaceo, come avviene negli atti amministrativi di stampo tradizionale"⁴³.

In altro caso si può richiamare la recente sentenza con cui il Consiglio di Stato ha confermato una decisione del giudice di primo grado modificandone la motivazione.

Mentre il giudice di primo grado aveva ritenuto che la decisione robotica non fosse legittima perché era stato completamente pretermesso il contributo dell'uomo, per il Consiglio di Stato il problema non è tanto questo⁴⁴ quanto piuttosto che la decisione algoritmica non è comprensibile nella sua motivazione e per questo va annullata.

Ed infatti per il Consiglio di Stato "l'utilizzo di procedure "robotizzate" non può, tuttavia, essere motivo di elusione dei principi che conformano il nostro ordinamento e che regolano lo svolgersi dell'attività amministrativa"⁴⁵.

Più di recente il giudice amministrativo di primo grado ha

⁴³ TAR, Lazio-Roma, sez. III bis, 22 marzo 2017, n. 3769.

⁴⁴ Per il giudice di appello "l'assenza di intervento umano in un'attività di mera classificazione automatica di istanze numerose, secondo regole predefinite (che sono, queste sì, elaborate dall'uomo), e l'affidamento di tale attività a un efficiente elaboratore elettronico appaiono come doverose declinazioni dell'art. 97 Cost. coerenti con l'attuale evoluzione tecnologica".

⁴⁵ Per il giudice di Palazzo Spada "la regola tecnica che governa ciascun algoritmo resta pur sempre una regola amministrativa generale, costruita dall'uomo e non dalla macchina, per essere poi (solo) applicata da quest'ultima, anche se ciò avviene in via esclusiva. Questa regola algoritmica, quindi: possiede una piena valenza giuridica e amministrativa, anche se viene declinata in forma matematica, e come tale, come si è detto, deve soggiacere ai principi generali dell'attività amministrativa, quali quelli di pubblicità e trasparenza (art. 1 l. 241/90), di ragionevolezza, di proporzionalità, etc.; non può lasciare spazi applicativi discrezionali (di cui l'elaboratore elettronico è privo), ma deve prevedere con ragionevolezza una soluzione definita per tutti i casi possibili, anche i più improbabili (e ciò la rende in parte diversa da molte regole amministrative generali); la discrezionalità amministrativa, se senz'altro non può essere demandata al software, è quindi da rintracciarsi al momento dell'elaborazione dello strumento digitale; vede sempre la necessità che sia l'amministrazione a compiere un ruolo ex ante di mediazione e composizione di interessi, anche per mezzo di costanti test, aggiornamenti e modalità di perfezionamento dell'algoritmo (soprattutto nel caso di apprendimento progressivo e di deep learning)"; deve contemplare la possibilità che - come è stato autorevolmente affermato - sia il giudice a "dover svolgere, per la prima volta sul piano 'umano', valutazioni e accertamenti fatti direttamente in via automatica", con la conseguenza che la decisione robotizzata "impone al giudice di valutare la correttezza del processo automatizzato in tutte le sue componenti". In definitiva, dun-

confermato l'orientamento per il quale la decisione robotica va dichiarata illegittima quando non se ne può ricostruire la motivazione ritenendo che: “ad essere inoltre vulnerato non è solo il canone di trasparenza e di partecipazione procedimentale, ma anche l'obbligo di motivazione delle decisioni amministrative, con il risultato di una frustrazione anche delle correlate garanzie processuali che declinano sul versante del diritto di azione e difesa in giudizio di cui all'art. 24 Cost., diritto che risulta compromesso tutte le volte in cui l'assenza della motivazione non permette inizialmente all'interessato e successivamente, su impulso di questi, al Giudice, di percepire l'iter logico – giuridico seguito dall'amministrazione per giungere ad un determinato approdo provvedimento”⁴⁶.

Conseguentemente la caratterizzazione “multidisciplinare” dell'algoritmo (costruzione che certo non richiede solo competenze giuridiche, ma tecniche, informatiche, statistiche, amministrative) non esime dalla necessità che la “formula tecnica”, che di fatto rappresenta l'algoritmo, sia corredata da spiegazioni che la traducano nella “regola giuridica” ad essa sottesa e che la rendano leggibile e comprensibile, sia per i cittadini che per il giudice.

La regola algoritmica, quindi, deve essere soggetta alla piena cognizione, e al pieno sindacato, del giudice amministrativo e tale esigenza risponde alla necessità di poter sindacare come il potere sia stato concretamente esercitato “ponendosi in ultima analisi come declinazione diretta del diritto di difesa del cittadino, al quale non può essere precluso di conoscere le modalità (anche se automatizzate) con le quali è stata in concreto assunta una decisione destinata a ripercuotersi sulla sua sfera giuridica”⁴⁷.

que, l'algoritmo, ossia il software, deve essere considerato a tutti gli effetti come un “atto amministrativo informatico”.

⁴⁶ TAR Lazio, Roma, III bis, 10 settembre 2019, n. 10964 (che a sua volta richiama TAR Lazio, Roma, n. 9924/2018).

⁴⁷ “Solo in questo modo – sempre per il Consiglio di Stato – è possibile svolgere, anche in sede giurisdizionale, una valutazione piena della legittimità della decisione; valutazione che, anche se si è al cospetto di una scelta assunta attraverso una procedura informatica, non può che essere effettiva e di portata analoga a quella che il giudice esercita sull'esercizio del potere con modalità tradizionali. In questo senso, la decisione amministrativa automatizzata impone al giudice di valutare in primo luogo la correttezza del processo informatico in tutte le sue componenti: dalla sua costruzione, all'inserimento dei dati, alla loro validità, alla loro gestione. Da qui, come si è detto, si conferma la neces-

Il giudice deve quindi poter sindacare la stessa logicità e ragionevolezza della decisione amministrativa robotizzata, ovvero della "regola" che governa l'algoritmo, di cui si è ampiamente detto.

7. La "povertà" di risorse dell'amministrazione italiana di fronte agli algoritmi.

Una delle tentazioni riscontrabili nel diritto positivo degli ultimi decenni è quella di voler realizzare riforme applicando la cd. clausola di invarianza finanziaria.

Ci si dimentica che le riforme camminano sulle gambe delle persone o più tecnicamente delle istituzioni, che per realizzarle occorre "pagare qualcosa" e ci si illude che l'aver inserito una prescrizione in un testo normativo possa, quasi miracolosamente, conseguire il risultato auspicato.

Mentre negli USA quando si è deciso di attuare il FOIA (Freedom of information act) si sono assunti circa quattromila funzionari a tempo pieno per una spesa annuale 500 milioni di dollari⁴⁸, nel nostro paese si è deciso di attuare il principio di trasparenza senza investire alcun tipo di risorsa. Quando in Cina si è deciso di attuare una seria politica ambientale si sono assunti a Pechino settantamila educatori ambientali, da noi si decide di attuare i principi dell'economia circolare a costo zero.

In altre parole l'aspetto organizzativo, ossia ciò che attiene alle strutture, agli uomini che realizzeranno le riforme non appare un aspetto secondario ma anzi ciò che maggiormente deve essere valutato.

E la riprova del fatto che ciò non sia avvenuto in materia di digitalizzazione della p.a. lo conferma il fatto che dal punto di vista organizzativo il governo del settore non abbia trovato una chiara, e soprattutto stabile, sistemazione⁴⁹: prima di chie-

sità di assicurare che quel processo, a livello amministrativo, avvenga in maniera trasparente, attraverso la conoscibilità dei dati immessi e dell'algoritmo medesimo".

⁴⁸ L. CARBONARA, *La difficoltà di decidere*, in *Giorn.dir.amm.*, 2018, p. 7.

⁴⁹ Sia sufficiente far riferimento al continuo mutamento di denominazione, funzioni e collocamento istituzionale dell'autorità di riferimento: dall'AIPDA, istituita nel 1993, al CNIPA, collocato presso la Presidenza del Consiglio dei Ministri, del 2003, a DigitPA del 2009, all'attuale Agid, Agenzia per l'Italia digitale, istituita nel 2012. La stessa natura giuridica è stata a lungo discussa: l'AIPDA era immaginata come un ente a metà tra l'autorità indipendente e

dersi cosa fare ci si sarebbe dovuti porre il problema di chi lo fa e con quali fondi.

Ecco dunque perché per fare una “predizione” sulle possibilità di sviluppo della decisione robotica si ritiene si debba far riferimento ad alcuni dati di contesto riferiti alla realtà dell’amministrazione italiana per lo meno per come si presenta nel 2019.

Un primo luogo comune da sfatare, al riguardo, è che i dipendenti pubblici in Italia siano troppi: si tratta di errore dal momento che il numero dei dipendenti pubblici in Italia, di circa 3.200.000 (più di un terzo nella scuola e una gran parte nella sanità)⁵⁰ è in valori assoluti considerevolmente più basso di quello di altri paesi a noi paragonabili come la Francia e la Germania (rispettivamente circa 5.500.000⁵¹ e 6.500.000⁵²).

Ma se anche lo rapportiamo alla popolazione dei tre paesi abbiamo i seguenti valori percentuali in Italia 5 persone su 100 sono impiegate nella p.a.; mentre in Francia e Germania 8 persone su 100 sono impiegate nella p.a.: questo significa che per portare i numeri dei nostri dipendenti pubblici allo stesso livello di quello di questi due paesi dovrebbero essere assunte più di 1.500.000 di persone ossia i dipendenti pubblici dovrebbero crescere di più del 50% (si noti che tra i dipendenti nel nostro paese il numero dei dipendenti dei Ministeri è di circa 145.000 persone⁵³).

Ecco dunque la prima sorpresa: gli impiegati pubblici in Italia sono pochi e questo è il primo dato che va tenuto presente quando si parla di riforme amministrative complesse come quelle del passaggio all’intelligenza artificiale nella decisione amministrativa che richiedono notevoli investimenti di risorse.

Ma vi è anche un secondo dato, in qualche modo ancora più rilevante del primo, quello per il quale l’età media dei dipendenti pubblici nel nostro paese è di cinquantadue anni.

un’agenzia, ma l’indipendenza è stata poi minata dalla collocazione del CNIPA presso la Presidenza del Consiglio, e all’attuale istituzione ora è riconosciuta esplicitamente, e dal punto di vista sistematico più correttamente, la natura di agenzia governativa, quindi sottoposta ai principi del d.lgs. 30.7.1999, n. 300: cfr. F. COSTANTINO, *Informatizzazione della p.a.*, cit., 2015.

⁵⁰ Fonte: Conto Generale della Ragioneria dello Stato, 2018.

⁵¹ Fonte: *Eurostat Statistic Explained, Archive: Public employment – France*, aggiornato al 2018.

⁵² Fonte: *Eurostat Statistic Explained, Archive: Public employment – Germany*, aggiornato al 2016.

⁵³ Dato rilevato al 24 aprile 2019: Fonte *Eurostat*.

Mettendo a confronto tale dato con quello relativo agli impiegati di altri paesi si vede innanzitutto come la nostra amministrazione sia la più vecchia in Europa e questo non fa ben sperare pensando che per affrontare queste tematiche di carattere innovativo occorrono risorse dalle menti giovani e elastiche⁵⁴.

Ciò significa che con gli attuali trend di assunzione nei pubblici concorsi (che si sono ridotti enormemente per motivi di bilancio), se si continua come ora, nell'arco di una quindicina di anni dell'amministrazione italiana non resterà più nulla.

Questi dati dovrebbero essere sempre tenuti presenti, come sullo sfondo, nel descrivere il perché, le motivazioni di fondo delle riforme a cui stiamo assistendo negli ultimi anni come quello delle liberalizzazioni (ad es. il *favor* per l'uso della SCIA) o dello sviluppo dell'uso del silenzio assenso senza dimenticare ovviamente che un'altra importante concausa è costituita dal fatto che le amministrazioni, per quello che si definisce il fenomeno dell'amministrazione difensiva⁵⁵, tendono a favorire tutte quelle riforme che eliminano la decisione o la discrezionalità.

In un contesto infatti in cui i dipendenti pubblici nel nostro paese tendono a non decidere per non assumersi responsabilità di fronte alla procura sia della Repubblica che della Corte dei Conti si crede che l'utilizzo di macchine che sostituiscano e an-

⁵⁴ Purtroppo, invece, nel nostro paese le questioni che riguardano la digitalizzazione vengono affrontate da generazioni che si sono formate in contesti tecnologicamente completamente diversi (da piccoli non avevamo i telefonini, internet) e che tendono ad affacciarsi alla pensione.

⁵⁵ Il termine amministrazione difensiva viene dal settore medico in cui si definisce medicina difensiva quella per la quale il medico per non incorrere in responsabilità tende a modificare la sua azione in senso positivo (ad esempio prescrivendo analisi, come radiografie, che non sarebbero strettamente necessarie e che possono danneggiare la salute del paziente oltre che comportare extracosti per l'erario) o in senso negativo (ad esempio rifiutandosi di intervenire in casi in cui le percentuali di successo siano basse). Uno degli episodi più antichi risale al IV secolo aC ed è riportato da Curzio Rufo nelle *Historie Alexandri Magni*: l'autore narra che Alessandro, gravemente ferito in battaglia, non riuscì a trovare alcun medico disponibile ad intervenire per asportare la freccia che si era conficcata nel suo corpo, sino a quando lo stesso, conscio della gravità della lesione e delle ragioni per le quali i chirurghi erano tanto restii a intervenire, promise saggiamente l'impunità a tal Critobulo, che infine lo operò: G. GUERRA, *La 'medicina difensiva': fenomeno moderno dalle radici antiche*, in *Salute e diritto*, 2013. Tra gli autori che se ne sono occupati in campo amministrativo G. CORSO, *Sfiducia, fiducia e giurisdizione*, in *Giorn. dir. amm.*, 2, 2017.

nullino la loro discrezionalità e la loro responsabilità possa essere accettato di buon grado.

Si può allora concludere che questa “povertà” di risorse amministrative crei spinte e contropinte in ordine all’uso della decisione robotica o del *government by data*: da una parte l’età avanzata dei dipendenti pubblici sembra essere un freno per lo sviluppo dell’applicazione dell’analisi predittiva all’attività amministrativa; dall’altra parte la necessità di rispettare i vincoli di bilancio e l’amministrazione difensiva potrebbero dare impulso allo sviluppo dell’amministrazione robotica.

Se prevarranno le spinte o le contropinte e se dunque si andrà o meno nel senso della decisione robotica sarà solo il tempo a dircelo: ma non si crede che dovremo aspettare molto per saperlo data la rapidità con cui si sono verificati sinora i mutamenti.

LA JUSTICE PRÉDICTIVE

Louis Larret-Chahine

SOMMAIRE: 1. Introduction. La réalisation d'une utopie. – 2. Définition de la justice prédictive. – 3. Les bénéfices de la justice prédictive. – A. Un outil d'aide à la décision pour les professionnels du droit. – B. Un levier de transparence et de performance du service public de la justice. – 4. Pour une critique constructive de la justice prédictive. – A. La nécessité d'un choix éthique. – B. La question de la régulation. – 5. Conclusion. Pour une université #legalgeek.

1. Introduction - La réalisation d'une utopie.

Le mouvement a commencé il y a soixante ou soixante-dix ans. Ce fut d'abord presque imperceptible, à la manière d'un lent glissement de l'horizon sur un navire. Plus tard, les historiens du droit retiendront peut-être une date encore antérieure, signal faible marquant les prémices d'un changement de paradigme judiciaire: celui de la fin de l'incertitude de l'issue des procès, aléa qui faisait le sel d'un contentieux mais l'angoisse des justiciables.

Les premiers à avoir formulé l'idée de l'anticipation d'un résultat contentieux en s'appuyant sur des décisions de justice antérieures s'appelaient Loevinger¹, Kort² ou Lawlor³. Avant cela même, d'autres avaient tenté de faire émerger la notion

¹ L. LOEVINGER proposait la création d'une nouvelle science «jurimétrique» définie comme «l'application de la méthode scientifique à l'étude du Droit», dont le but était le «calcul des probabilités des décisions judiciaires», *Jurimetrics The Next Step Forward*, dans *Minnesota Law Journal*, 1949.

² F. KORT, *Predicting Supreme Court Decisions Decisions Mathematically: A Quantitative Analysis of the «Right to Counsel» Cases*, dans *The American Political Science Review*, mars 1957, n°51(1), p. 1.

³ R. C. KIRCHNER LAWLOR, *What Computers Can Do: Analysis and*

de probabilité des jugements⁴. Comme l'explique Guillaume Zambrano, le concept de prédiction jurisprudentielle en lui-même «a été proposé il y a plus d'un siècle déjà par le réaliste américain Holmes – professeur à Harvard puis juge à la Cour Suprême – lors d'une conférence prononcée à Boston en 1987 (...): "ce que j'entends par Droit ce sont les prophéties de ce qui sera concrètement décidé par les tribunaux"»⁵.

Ces travaux, pour la plupart, ont été très mal accueillis, jugés coupables d'une approche considérée incompatible avec cette science humaine qu'est le droit. Ainsi, alors que la plupart des secteurs s'appropriaient les outils technologiques pour entamer leur mue, les juristes ont adopté un comportement insulaire.

Malgré tout, début 2017, sous l'effet conjugué des progrès informatiques et de la constitution de bases de données judiciaire, la justice prédictive est devenue une réalité⁶. La pensée de Holmes, pour qui la science du droit devrait être une science prédictive serait-elle en train de s'imposer ?

Pour Antoine Garapon, «les défis que la justice prédictive lance à l'idée de justice» entraîneraient une «révolution cognitive»⁷, qui irait jusqu'à modifier la définition, ou au moins la perception de la justice et du droit.

Cette dynamique confirmerait ainsi une idée séduisante, dérivée de la théorie de l'information formulée par Claude Shannon en 1948⁸, selon laquelle finit toujours par s'imposer le meilleur vecteur de contenu. Même si cela prend parfois du temps (le papier face au papyrus, internet sur le minitel,

Prediction of judicial Decision, dans *American Bar Association Journal*, avril 1963, n°49, p. 337.

⁴ Simon-Denis Poisson publie ses «Recherches sur la probabilité des jugements en matière criminelle et en matière civile» en 1837; Nicolas Bernoulli soutient une thèse sur l'utilisation judiciaire du calcul des probabilités en 1709, sobrement intitulé «De Usus Artis Conjectandi in jure».

⁵ G. ZAMBRANO, *Précédents et prédictions jurisprudentielles à l'ère des big data: parier sur le résultat (probable) d'un procès*, 2015, hal-01496098.

⁶ La commercialisation du logiciel développé par la legaltech Predictice (www.predictice.com) a été lancé en France le 8 juin 2017 en partenariat avec le groupe Wolters Kluwer. (<https://news.predictice.com/wolters-kluwer-et-predictice-sallient-pour-d%C3%A9ployer-la-justice-pr%C3%A9dictive-en-france-e24243517919>)

⁷ A. GARAPON, *Les enjeux de la justice prédictive*, dans *Revue pratique de la prospective et de l'innovation*, octobre 2016.

⁸ C. E. SHANNON, *A Mathematical Theory of Communication*, dans *Bell System Technical Journal*, vol. 27, 1948, pages 379-423 et 623-656.

l'email contre le fax, les systèmes prédictifs sur les bases de données brutes...), la prime irait *in fine* au système capable de transporter la quantité maximale d'information. Cependant les algorithmes de traitement, comme ceux de compréhension du langage naturel, utilisés pour faire fonctionner un logiciel de justice prédictive imposent une logique encore un peu différente.

Car ils analysent, compilent puis synthétisent une immense quantité d'information, ils sont certes un vecteur de contenu, mais aussi un *créateur* de contenu. Les algorithmes sont devenus un médium. La justice prédictive est donc à la fois une nouvelle manière de créer, de transférer mais aussi de comprendre l'information.

2. Définition de la justice prédictive.

La justice prédictive est la projection dans le futur d'une statistique réalisée sur un panel de décisions de justice passées similaire au cas en question.

Sous un angle fonctionnel, les algorithmes de justice prédictive permettent de calculer les chances de succès d'un procès devant une juridiction, le montant des indemnités qu'il est possible d'obtenir, et d'identifier les arguments de fait et de droit qui auront le plus d'impact sur seront les plus susceptibles d'influer sur la décision juridictionnelle à venir.

Si les pouvoirs publics, une partie de la doctrine et des professionnels du droit soutiennent l'émergence de la justice prédictive, c'est d'abord en raison des immenses bénéfices qui en sont attendus (3). Ces espoirs ne doivent cependant pas empêcher une critique constructive du système en train de s'imposer (4).

3. Les bénéfices de la justice prédictive.

Face à cette révolution annoncée comme majeure⁹, mais

⁹ La justice prédictive va permettre d'atteindre «un niveau de réalité qui était jusqu'à présent inaccessible. C'est non seulement toute la production qui devient transparente, mais aussi le détail, juge par juge, argument par argument, partie par partie. (...). La justice prédictive bouleverse aussi la

provoquée par un programme informatique dont il est difficile pour ses utilisateurs de comprendre le mécanisme, qu'aucun juriste ne peut monter et démonter entre ses mains pour comprendre de quelle manière il est constitué, les professionnels du droit balancent entre deux attitudes.

Certains, précautionneux ou conservateurs, partent du principe que le progrès technologique ne peut que perturber une discipline vieille de trois millénaires¹⁰. L'attitude à adopter vis-à-vis de la prédiction des jugements serait donc la même que celle qu'il fallait avoir vis-à-vis du fax, du téléphone, ou d'internet: au mieux la circonspection, au pire le rejet¹¹.

D'autres, peut être conscients que la technologie a en réalité toujours bénéficié aux professionnels du droit (traitement de texte, dématérialisation des procédures, base documentaire en ligne, email...) ou plus opportunistes¹² adoptent une attitude positive. Pour eux, le monde se numériserait et se judiciaireiserait en même temps: face à un besoin de sécurité juridique accru, il faut donc s'adapter aux évolutions du marché et se doter des outils les plus performant.

Tous ont par contre le même sentiment d'accélération, comme si la digue patiemment construite durant les dernières décennies venait de céder face à la trop forte pression technologique accumulée à l'extérieure. C'est ce phénomène de rattrapage du secteur juridique, de comblement du décalage avec d'autres secteurs économiques, qui est à l'origine de cette tension parfois constatée¹³.

fonction ordonnatrice du temps car elle introduit *ab initio* si ce n'est la solution du moins une solution très probable. (...). La justice prédictive augmente le présent», A. GARAPON, op. cit.

¹⁰ Si l'on prend comme point de départ le Code d'Hammurabi, stèle juridique babylonienne datée d'environ 1750 av. J.-C.

¹¹ F. REED DICKERSON, *Automation and the lawyer*, dans *Articles by Maurer Faculty*, 1965, paper 1561: «the notion that someone might rely on a little black box to predict judicial behavior has created near hysteria in some member of the bar».

¹² «In fact, those lawyers who recognize the pending changes as an opportunity will likely do very well in this new environment», R. VOGL, *The coming of Age of Legal Technology*, dans *Stanford Law School*, 26 September 2016.

¹³ Un autre secteur, celui du journalisme, semble avoir vécu eu des réticences proches: «One strategy developed by both journalists and magistrates is to draw on their professional ethos to criticize the legitimacy of metrics and refuse to engage with them. They argue that "good justice"»

Néanmoins, n'en déplaise aux fantasmes liberticides de certains rétifs, «le probable n'est qu'un canton du possible»¹⁴; en d'autres termes l'avenir est dans le domaine de l'action et de la volonté, pas de l'informatique.

En effet, la justice prédictive ne vise pas à automatiser la justice, mais seulement à fournir un outil d'aide à la décision à destination des professionnels du droit (A), et à contribuer à la performance et la transparence du service public de la justice (B).

A. Un outil d'aide à la décision pour les professionnels du droit.

Anticiper est un exercice naturel. Intuitivement, beaucoup de juristes (et en particulier les avocats) utilisent déjà le raisonnement prédictif en se basant sur trois catégories d'information: le droit, les caractéristiques premières du litige et les éléments de contexte qu'ils apprécient en fonction de leur expérience.

L'exemple du calcul d'indemnités dans le cas d'un licenciement est probant. Il existe des règles précises pour déterminer les indemnités d'un salarié (le droit), qui s'appuient sur des éléments factuels et concrets tels que l'ancienneté dans l'entreprise (les caractéristiques du litige). Mais les indemnités totales peuvent varier, par exemple selon l'état de santé physique ou moral du salarié (éléments de contexte).

La justice prédictive n'a pas pour vocation de remplacer l'expert dans son travail de compréhension et d'analyse, mais de lui permettre de mieux saisir l'impact de ces éléments de contexte sur la décision finale. Il s'agit d'ajouter à son expérience le résultat de l'analyse d'un nombre considérable de données (le fameux *big data*), infiniment plus que ce que le cerveau humain est capable de compiler et de traiter.

Ainsi, la technologie de justice prédictive est avant tout une aide à la décision destinée aux professionnels du droit,

or "good journalism" cannot be measured and evaluated in quantitative ways", A. CHRISTIN, *The Hidden Story of How Metrics Are Being Used in Courtrooms and Newsrooms to Make More Decisions*, dans *Ethnography Matters, Co-Designing With Machines, Special Issue*, 2016.

¹⁴ P. MASSÉ, *De la prospective, Textes fondamentaux de la prospective française, 1955-1966*, 2eme édition, 2007, Paris.

permettant d'ajuster intelligemment conseils et stratégie en dépassant la connaissance empirique et nécessairement imparfaite du praticien.

Comment cela fonctionne-t-il ?

Pour faire simple, l'analyse prédictive se décompose en trois étapes, qui reviennent à répondre à trois questions.

1 – «Que s'est-il passé ?»: grâce à l'enrichissement des données, c'est-à-dire à l'ajout de métadonnées décrivant les caractéristiques du litige, cette solution est d'abord un formidable moteur de recherche, sémantique et intelligent, permettant de proposer au professionnel toutes les informations (textes, jurisprudences, informations diverses) dont il aura besoin.

2 – «Que va-t-il se passer ?»: puisqu'il n'existe jamais deux litiges parfaitement identiques, l'enjeu est alors de comprendre l'impact d'un facteur ou d'une combinaison de facteurs (analyse multivariée) sur la résolution. C'est ici qu'entrent en jeu les algorithmes de *machine learning*, capables de croiser les observations pour créer des modèles prédictifs complexes. En appliquant ces modèles aux caractéristiques d'un litige, il est ainsi possible d'évaluer les probabilités de résolution.

3 – «Comment l'optimiser ?»: la technologie est capable d'évaluer et de comparer plusieurs stratégies contentieuses, habilitant l'utilisateur à choisir l'option qui a statistiquement le plus de chances de succès, en fonction des caractéristiques variables de l'affaire.

Les professionnels du droit, en se dotant de cet outil de performance, peuvent ainsi se concentrer sur le cœur de leur métier et leur valeur ajoutée¹⁵.

¹⁵ En ce sens, la décision de s'équiper de ce type de logiciel est similaire à celle d'une clinique achetant un scanner dernier cri, renforçant par exemple la valeur de l'avocat dans l'élaboration de la stratégie et son expertise auprès de la clientèle, comme l'explique le Bâtonnier de l'Ordre des avocats de Lille, Stéphane Dhonte, dans une interview réalisée par L. GARNERIE, *La justice prédictive ne tuera pas le métier d'avocat*, dans *Gazette du Palais*, p. 9, février 2017.

B. Un levier de transparence et de performance du service public de la justice

1 – Un barème intelligent et évolutif

La qualité du service public de la justice devrait s'améliorer au fur et à mesure du déploiement de la justice prédictive au sein de la magistrature. En effet, cette technologie permet d'éviter l'uniformisation que peut induire un système de barème (système qui tente une forme d'anticipation, mais en lissant les cas particuliers au préjudice de la variété du réel et de l'expérience des juges). La catégorisation par barème n'étant jamais assez fine, elle se fait, en réalité, au détriment du justiciable.

A l'inverse, un calcul statistique, avec une pondération fine des différents critères, s'appuyant sur l'ensemble des précédents permet une meilleure prise en compte du particularisme. La justice prédictive réussit donc à répondre à la triple exigence d'anticipation des coûts pour une entreprise, de transparence du système judiciaire, et de la prise en compte des particularités. En d'autres termes, cela permet d'éviter une barémisation brutale pour s'orienter vers un système modérateur, qui homogénéise la façon dont le droit est rendu sur le territoire en s'appuyant sur l'intelligence collective des magistrats, tout en informant sur l'application concrète de la règle de droit au sein des juridictions. La justice deviendrait ainsi plus éthique, s'inspirant de la règle de plomb des architectes de Lesbos qui s'adaptait à ce qu'elle mesure. Grâce à la technologie prédictive, la justice pourrait devenir une balance de cristal.

2 – La médiation: une externalité positive des algorithmes de justice prédictive ?

Savoir ce qu'il se passera probablement en cas d'action contentieuse incite à l'utilisation des modes alternatifs de règlement des litiges.

Par exemple, commencer une négociation avec la probabilité de gagner ou de perdre à 90% un montant de 100.000 euros pousse, dans la plupart des cas, à transiger à hauteur de 90.000 euros, s'évitant par là-même plusieurs années dans les prétoires et des coûts de procédure élevée.

La justice prédictive remplit dans certains cas une fonction d'épouvantail. La médiation, qui permet de préserver des relations commerciales, familiales ou personnelles, s'en trouve

facilité. C'est ce qu'ont compris et constaté les assureurs de protection juridique, dont plusieurs utilisent la solution Predictice pour augmenter significativement leur taux de déjudiciarisation.

La technologie de justice prédictive devrait donc entraîner progressivement un désengorgement des juridictions¹⁶ par la disparition des cas les moins sujet à variation. Par un effet de vase communicant, la médiation, la transaction et les autres règlements amiables des conflits vont se multiplier, développant un nouveau système économique de résolution des litiges, probablement intégralement dématérialisé, rapide, et peu onéreux.

Ces effets bénéfiques, pour les professionnels comme pour les justiciables, ne doivent cependant pas empêcher l'émergence d'une critique constructive de la justice prédictive (3).

4. Pour une critique constructive de la justice prédictive.

La question de la régulation des algorithmes (B), qui est trop souvent perçue comme la seule réaction possible, ne peut pas se passer d'une réflexion et d'un choix éthique sur la position à adopter vis-à-vis la technologie de justice prédictive (A).

A. La nécessité d'un choix éthique.

L'ethnographe Tricia Wang le rappelle: l'opposition manichéenne du monde des hommes et de celui des machines n'a pas de sens¹⁷. En effet, les progrès des algorithmes sont le résultat d'interactions croissantes entre les deux univers, d'abord au moment de la construction, puis de «l'entraînement», et enfin du perfectionnement.

¹⁶ Le stock d'affaire des Cours d'appel, Tribunaux de grande instance et des conseils de prud'hommes (hors référé) était de 1.239.000 affaires au 31 décembre 2015, Ministère de la Justice/SG/SDSE, Répertoire général civil, rapport annuel de la Cour de cassation.

¹⁷ Notamment dans son post sur le blog Ethnography Matters (<http://ethnographymatters.net/blog/2016/06/13/co-designing-with-machines-moving-beyond-the-humanmachine-binary/>)

Les algorithmes peuvent se tromper. Ils peuvent être injustes. Ils peuvent pérenniser un état de fait non désiré. Il arrive également que les algorithmes répercutent les choix de leurs créateurs, néo-scribes dont les productions ont de plus en plus de conséquence sur le monde non-virtuel – monde dont la limite se dissout devant nos yeux.

L'enjeu éthique doit donc être sérieusement adressé. La réflexion que mène le Comité éthique et scientifique de la justice prédictive, hébergé par le cabinet Taylor Wessing¹⁸ (qui a d'ailleurs obtenu un accès aux travaux de la startup Predictice, leader sur le secteur en France) devenait d'ailleurs urgente avec le début de l'utilisation de cette solution par des magistrats français dans deux cours d'appel¹⁹.

Comme a pu l'exprimer le professeur Christophe Jamin, Directeur de l'École de Droit de Sciences Po et membre du comité, «la révolution engendrée par Predictice est telle qu'on ne peut s'abstenir de réfléchir collectivement à toutes les conséquences qu'elle induit, dont certaines ne sont pas encore envisagées».

Le premier des risques est celui de l'effet performatif: l'algorithme aurait un rôle de prophétie auto-réalisatrice, tordant le réel pour le faire correspondre à son anticipation. Cet effet secondaire a néanmoins surtout été constaté dans les systèmes complètement automatisés, dans lesquels la recommandation et la décision sont prises par une machine. Le risque de distorsion et d'homogénéisation serait donc plus faible dans la justice, principalement car les interventions humaines sont nombreuses.

Il semble toutefois nécessaire de pouvoir mesurer cet effet performatif. Plusieurs projets de recherche sont d'ailleurs en cours, soit dans des universités pour analyser l'impact de la justice prédictive à l'occasion de simulation de procès, soit directement dans les zones où le déploiement de Predictice est le plus avancé.

Un deuxième risque existe avec l'application de ces outils

¹⁸ <https://france.taylorwessing.com/fr/la-premiere-reunion-du-comite-ethique-et-scientifique-de-la-justice-predictive-s-est-tenue-dans-les-locaux-du-cabinet-taylor-wessing>

¹⁹ <http://www.lefigaro.fr/actualite-france/2017/05/05/01016-20170505-ARTFIG00-260-les-juges-experimentent-la-justice-predictive.php>

d'aide à la décision en matière pénale, réveillant le fantasme dickien d'un monde où les rapports minoritaires²⁰ (le futur improbable, les quelques pourcents alternatifs, l'incertitude en somme) sont effacés et permettent de condamner avant même la commission du délit ou du crime. Bien que le parti pris soit aujourd'hui de ne pas faire fonctionner la technologie de justice prédictive en matière pénale, il faut là-aussi rester vigilant aux effets pervers, en particulier en matière de récidive²¹, ou comme cela a été souligné dans un domaine aux enjeux proches: la police prédictive²².

Dans l'attente du résultat des premières études qui devraient aider à sortir des postures de principe sur les enjeux éthiques, il importe de garder en tête que cette technologie constitue avant tout une aide à la décision, qu'elle a ses équivalents dans d'autres domaines, et que ses évolutions sont suivies de près par des professionnels attentifs. Il importe donc de ne pas freiner une innovation aux multiples bénéfices par une entreprise de régulation mal calibrée.

B. La question de la régulation.

La régulation, bien qu'elle soit «synonyme d'équilibre entre les pouvoirs et de reconstruction des rapports de force»²³ s'est réduite, s'agissant de la justice prédictive, au prisme de la régulation par la mise à disposition des données publiques fournies par l'Etat, ici unique producteur des données nécessaires.

Le principe de l'accès aux décisions de justice n'est pas contesté dans son essence. La loi pour une République numérique a en effet posé le principe de la mise à disposition,

²⁰ Le titre original de la nouvelle de P. K. DICK est *The minority report*, publié en 1956 dans la revue *Fantastic Universe*.

²¹ J. ANGWIN, J. LARSON, S. MATTU, L. KIRCHNER, *Machine Bias, There's software used across the country to predict future criminals. And it's biased against blacks*, dans *ProPublica*, mai 2016.

²² «La critique de Tim Hope permet de rendre compte de l'ignorance qui évolue avec le développement des connaissances sur la prédiction des victimisations et des innovations qui leur sont associées», B. BENBOUZID, *De la prévention situationnelle au predictive policing*, dans *Champ pénal Vol. XII*, 2015.

²³ Définition du droit de la régulation économique, par M.-A. FRISON-ROCHE, dans *Receuil Dalloz*, 2004.

à titre gratuit et dans un format réutilisable (sous réserve du respect de la vie privée des personnes concernées dans les décisions) de l'ensemble des décisions de justice²⁴, fournissant la matière première nécessaire au fonctionnement des algorithmes.

Par contre, celle des modalités de la mise à disposition n'est pas réglée, en particulier sur le point sensible du degré d'anonymisation des décisions.

L'objectif principal de l'anonymisation est d'empêcher la ré-identification des personnes physiques concernées par les décisions de justice et ainsi préserver leur vie privée. Ce souci se heurte rapidement à un défi technique. D'un côté, si l'on enlève systématiquement tous les éléments caractéristiques (ancienneté d'un salarié, lieu de travail, montant du salaire, motif du licenciement, participation aux instances représentatives, etc.), ce qui est parfois nécessaire pour garantir l'anonymat, alors la lecture de la décision ou son analyse perdra tout son intérêt. De l'autre, si on laisse ne serait-ce que quelques éléments, avec l'évolution des techniques d'identification dynamique qui croisent de multiple base de données, alors l'anonymisation sera illusoire.

La voie la plus simple pour assurer l'application de la loi est pourtant évidente: conserver un système de pseudo-anonymisation (en remplaçant les noms propres par X ou Y) et appliquer le système répréhensif prévu par l'article 226-19 du code pénal²⁵. Si cela est insuffisant, il sera toujours

²⁴ Voir en particulier les articles 3, 20 et 21 de la loi n°2016-1321 du 7 octobre 2016. L'Union Européenne semble également favorable à ce mouvement de transparence (la directive 2003/98/CE du 27 novembre 2003 consacre le principe de réutilisation des informations publiques), dont il est possible de trouver une justification jusque dans l'article 15 de la Déclaration des droits de l'Homme et du citoyen, selon lequel «La Société a le droit de demander compte à tout agent public de son administration». Le Parlement européen («Greater access to court files for third persons is not only recommended, it is necessary in view of the abovementioned problems ranging from some inconveniences to infringements of procedural rights, acknowledged as a fundamental human rights», “National practices with regard to the accessibility of court document”, study, Policy department, 2013) comme la Cour Européenne des Droits de l'Homme (Werner v. Austria, Szucs v. Austria, Sunday Times v. United Kingdom 1979) sont également de cet avis.

²⁵ «Est puni de cinq ans d'emprisonnement et de 300 000 euros d'amende (...) le fait, hors les cas prévus par la loi, de mettre ou de conserver en mémoire

possible de créer un délit de ré-identification, ou d'alerter la Commission Nationale de l'Informatique et des Libertés (CNIL), première autorité administrative indépendante, afin qu'elle se saisisse de la question, voir qu'elle l'intègre dans son programme de contrôle prioritaire.

L'enjeu de la mise à disposition des données est plus élevé qu'il n'y paraît à première vue. Les tentatives pour croiser le droit et les mathématiques sont nombreuses, notamment aux Etats-Unis et en Chine. Affirmer l'indépendance technologique de la France, favoriser son leadership et donc l'influence de son droit dans le monde est un combat économique crucial à moyen terme. Or, sans données, les possibilités de développement et d'entraînement des algorithmes sont beaucoup plus limitées.

5. Conclusion – Pour une université #legalgeek.

Le *big data* judiciaire est probablement en train de provoquer un glissement notionnel du terme «jurisprudence»: de la norme à vocation générale élaborée par les hautes juridictions, l'acception basculerait vers l'interprétation statistique donnée d'une problématique juridique par l'ensemble des décisions de justice qui y sont relatives²⁶. Dérivé du latin *pro-video* (je vois avant), le terme jurisprudence retrouverait ainsi, grâce à sa prévisibilité, son sens premier²⁷.

A court terme, l'étude de la jurisprudence devrait ainsi tendre vers une forme d'étude probabiliste capable de décrire l'application réelle de la loi, de croiser ces informations avec de multiples facteurs (économiques, sociologiques, politiques...) et ne se limitera plus à un discours d'opinion sur

informatisée des données à caractère personnel concernant des infractions, des condamnations ou des mesures de sûreté.

²⁶ A bien y regarder, cette définition n'est d'ailleurs pas si éloignée que celle qu'en donne G. CORNU: «la jurisprudence désigne une habitude de juger dans un certain sens ou une tendance habituelle d'une juridiction déterminée à juger dans tel sens», *Vocabulaire juridique*, sous la direction de A. H. CAPITANT, 9^e éd., 2011, Paris, p. 587.

²⁷ «La jurisprudence est l'aptitude à prévoir le sens d'une décision avant qu'elle ne soit rendue à partir de l'expérience des décisions antérieures et de l'analyse des motifs en logique juridique. La jurisprudence est une prédiction», G. ZAMBRANO, *op. cit.*

l'interprétation des normes. La révolution méthodologique souhaitée par les théoriciens réalistes du Droit s'accomplira donc par la technique.

En ce sens, si l'utilisation de la technologie de justice prédictive nécessite bien sur une solide connaissance juridique - notamment pour distinguer la corrélation de la causalité, et optimiser au mieux la stratégie judiciaire - elle implique aussi une évolution du travail d'analyse et des compétences à mobiliser pour le faire.

L'université et les lieux d'enseignement en ont parfaitement conscience et commencent d'ailleurs à s'adapter. Pour Elizabeth Menesguen, ancien Bâtonnier, ancien directeur de l'École de Formation des Barreaux de la Cour d'Appel de Paris et Président délégué de la Commission formation du Conseil National des Barreaux, «Predictice est un outil précieux pour les générations à venir pour mieux servir nos concitoyens». C'est la raison pour laquelle un partenariat a été conclu afin que l'ensemble des Ecoles de formation d'avocat puissent notamment bénéficier d'une intervention sur la justice prédictive. Ce Programme Predictice pour l'Enseignement et la Recherche (PPER)²⁸ a d'ailleurs été étendu aux universités.

Au-delà même de l'objectif crucial de la formation des juristes de demain, «la prédiction-connaissance, en permettant de donner une photographie, voire une cartographie, du contentieux, ouvre de nouvelle perspective de recherche. Jusqu'alors, une connaissance exhaustive de toutes les décisions dans un contentieux donné n'était pas possible»²⁹.

«C'est le fait qui fait le droit» affirmait Antoine Loysel (1536-1617). La justice prédictive pourrait bien lui donner raison.

²⁸ Le détail du Programme Predictice pour l'Enseignement et la Recherche est disponible ici: <https://news.predictice.com/le-programme-predictice-pour-lenseignement-et-la-recherche-5d28cddb794b>.

²⁹ *La justice prédictive: le contentieux dévoilé*, projet de recherche dans le cadre d'une candidature au GIP Justice de Morgan Sweeney, Maître de conférence à l'université Paris-Dauphine, 2017.

LAW AS A SET OF DECISIONS.
ON MERITS AND DANGERS OF LEGAL REALISM
THROUGH THE PRISM OF BIG DATA ¹

Wojciech Zagorski

TABLE OF CONTENTS: 1. Introduction. – 2. The Triumph of Legal Realism. – 3. The Dangers of Legal Realism.

1. Introduction.

“Justice is what the judge ate for breakfast.” This is how Jerome Frank, a well-known American judge, allegedly summarized the realist theory of law. The statement was obviously provocative, but the idea itself is more profound than one might think at first glance. Instead of defining law as a set of rules, a body of norms - as we usually do in Europe -, American lawyers insist on the voluntarist dimension of judicial decision-making. Rules, they say, are not the reason, but the product of judicial decisions. But rules do not decide in any given case. Judges do. When deciding on individual cases, judges are guided not only by legal rules, but also by a whole series of nonnormative factors: by some specificities of the case, by general tendencies and necessities of the time, by their own experience and personal preferences. Therefore - as realists say -, the law is basically not a set of objective rules, but a set of individual decisions, determined by much more than a mere precedent, or statute.

This way of thinking should gain in popularity among

¹ I should like to extend my thanks to Louis Larret-Chahine, CEO and co-founder of *Predictice*, a leading French company involved in big data exploration for legal purposes. All of the data presented hereby were extracted using *Predictice* software.

European lawyers, as we are becoming more and more familiar with big data analytics provided by new software tools, such as *Predictice*. For the first time in history, we will soon be able to access and to analyze all judicial decisions. Indeed, a recent French statute provides for a mandatory digital publication of all decisions made by the courts, meaning that some 3 million cases will be added every year to existing legal French databases². Combined with new software capacities, this will give us an unprecedented insight into the reality of the so-called *law in action*, especially with respect to lower courts' practice.

What strikes us the most when we examine data provided by these new software tools, is the huge outcome disparity between judges and between courts applying the same legal rules to similar cases. Data analytics tends to confirm the realist approach to the law, showing that the same legal rules are not necessarily applied in the same manner by different lawyers. To take only one example here, if the average amount of damages in cases involving wrongful termination of a job contract is two times higher in Metz, than what is usually awarded by judges in Orléans, or in Nancy, we should wonder if such a disparity could be justified by any kind of objective differences between cases. What those numbers show us, is that realists were at least partially right. There is more to law than simply rules and facts of the case. Personal approach, ideology, individual convictions and preferences are important factors of judicial decision-making. By confirming realist intuitions, big data helps us to recognize the voluntarist dimension of judicial decision.

Of course, the more suspicious we are of judicial bias, the more rigorous we become about procedural justice. If fair trial and due process guarantees are so well developed today, it is because, at some point of our common legal history, we became aware of the inherent subjectivity of judicial decision. Legal realism was an important part of that process. What big data gives us today, however, is scientific hard proof, and this is new. Until now - especially in Europe -, legal realism was merely a theory, a philosophical posture, a conviction fueled

² Loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice, Article 33 (JORF n°0071 du 24 mars 2019, texte n° 2).

by modern skepticism. While shared by many, realism was rejected by other lawyers. As we speak today, big data analytics enables us to assess objectively the extent to which judges actually bend rules. It would be difficult to predict the practical impact of those new statistic arguments in a long-term perspective, but the outcome will not be purely theoretical. While legal realism has already had a huge influence on American law, big data will surely leave its mark on European legal culture.

However, there might be a price to pay for the truth about judicial subjectivity. We may already observe a dangerous tendency within American legal culture towards excessive politicization of the judiciary. We all saw what happened during Senate confirmation hearings of Brett Kavanaugh, the U.S. Supreme Court justice. It was a somewhat unpleasant experience to watch a Supreme Court nominee answer questions about his sexual life in college, and one may wonder if it served well the authority of the U.S. judiciary. There is a direct link between those events and the triumph of legal realism. Indeed, if judicial decision is essentially a political one, we should treat judges in the exact same way we treat other politicians. In that case, instead of asking them about legal rules or precedents, it may seem normal to ask Supreme Court nominees about their personal convictions and experiences, as those factors effectively determine the outcome of cases presented to the SCOTUS. We should be aware of what it could mean for the very existence of independent judicial review. What is at stake here is our idea of separation of powers. If all decisions are intrinsically political, why should there be any space for the so-called independent judiciary? Why should we have judges appointed for a life tenure, immune from democratic accountability? Those questions will inevitably arise, as we become more and more aware of judicial bias.

What I would like to discuss briefly here, are two points. Firstly, I will try to show how big data analytics is leading towards the ultimate triumph of realist theory of law. Then, I will try to reflect on some risks related to such an outcome.

2. The Triumph of Legal Realism.

Figure 1 shows some interesting statistics related to the indemnities granted in France for unlawful dismissal of employees, based on 77 472 cases analyzed by *Predictice* software. We can observe significant disparities between different Appellate Courts. For instance, the average indemnity granted by Saint-Denis de la Réunion Court is almost 80% higher than the one granted by Bastia Court in Corsica. The outcome disparity could be related to revenue differences between two islands, but the average revenue *per capita* is actually quasi-identical in both cases³. Another possible explanation could be a geographical one. La Réunion is situated in the open Indian Ocean, whereas Corsica is situated in the Mediterranean. Therefore, la Réunion Court could be inclined to grant the maximal amount of the indemnity provided within the letter of the statute, as workers mobility in the Indian Ocean is severely restricted in comparison with Mediterranean islands. Therefore, an unlawfully dismissed employee would have more difficulties to find another job.

Fig. 1. Wrongful termination indemnities granted by French appellate courts (based on 77472 cases). Source: *Prédicite* (2019).

Circuit	Average indemnity amount	Number of indemnities granted	Total number of indemnity claims
France (all appellate courts combined)	€ 9 600	31 600	77 472
Saint-Denis de la Réunion	€ 10 200	440	967
Bastia (Corsica)	€ 5 700	121	248
Montpellier	€ 5 700	930	1916

³ According to French *National Institute of Statistics and Economic Studies (INSEE)*, the average revenue *per capita* in Bastia (in 2014) was € 2 233 net per month. In Saint-Denis de la Réunion, the capital city of la Réunion island, the average revenue was € 2 279 net per month (2014).

Some disparities, however, seem much more difficult to explain in any objective manner. As Figure 1 shows, the average indemnity granted for wrongful termination by French appellate courts is € 9 600, whereas the average indemnity allowed by Montpellier Appellate Court is only € 5 700, 40% less than the national average. In that case, the geographical explanation is simply not valid, as Montpellier is situated on the Continent. At the same time, the average revenue in Occitanie (Montpellier area) is actually slightly higher than the national average income (€ 2 385 net/month in Occitanie *versus* € 2 225 in France, according to *INSEE*, 2014). Finally, one might think that low indemnity rates may be due to some specific characteristics of cases presented to Montpellier Court. However, numbers cited above reflect the average outcome of 930 different cases. It would be difficult to assume that judges in Montpellier handled more than 900 cases radically different from what is usually presented to other appellate courts in France.

Mobbing cases judged by French appellate courts display yet another kind of disparity (Fig. 2). Generally speaking, the average success rate of an action for workplace harassment is 50%, which is not surprising by any means. Some exceptions to the rule, however, seem quite intriguing. For instance, in Grenoble, the average success rate of the same action is as high as 68%. At the same time, other appellate courts admit only one out of three claims of this type. In practice, those numbers mean that workers suing for mobbing in Grenoble have - statistically speaking - almost two times higher chance of winning the case, not to mention the average outcome in Bastia Appellate Court, where the success rate is only 13%. One can only wonder what could possibly explain such a difference, if not different judge's attitude towards mobbing cases.

Fig. 2. Average success rate. Mobbing cases judged by French appellate courts (based on 8391 cases). Source: *Prédicite* (2019).

Circuit	Average success rate	Cases won by plaintiffs	Total number of mobbing cases
France (all appellate courts combined)	50 %	4 253	8 391

Circuit	Average success rate	Cases won by plaintiffs	Total number of mobbing cases
Paris	52 %	782	1 505
Grenoble	68 %	242	355
Pau	37 %	57	154
Reims	38 %	50	132
Chambéry	37 %	47	126
Bastia	13 %	4	30

Finally, the comparison of average indemnity granted in mobbing cases reveals some interesting disparities too. As figure 3 shows, the average amount of damages allowed in Metz is twice as high as in Orléans or Nancy. There is no objective factor explaining this kind of variation. It would be difficult to assume that some case specifics may actually justify such a different outcome, as numbers below reflect a total of 80 sentences issued in Orléans, and more than 40 cases judged both in Metz and Nancy. Judges may be confronted with one or few exceptional cases, but 42 mobbing cases in Metz, after all, should not be much different from 42 mobbing cases in Nancy.

Fig. 3. Mobbing indemnities granted by French appellate courts (based on 8 391 cases). Source: *Predictice* (2019).

Circuit	Average indemnity amount	Number of indemnities granted	Total number of indemnity claims
France (all appellate courts combined)	€ 9 200	2 575	8 391
Basse-Terre (Guadeloupe)	€ 18 500	20	62
Metz	€ 12 600	42	116
Paris	€ 12 300	533	1 505
Orléans	€ 5 700	80	213
Nancy	€ 5 200	42	160

If those disparities may seem surprising, one should remember that what is compared here is the activity of appellate courts taken as a whole. In other words, the numbers above reflect the attitude of all judges working for a given court, which is a kind of local average on its own. Statistical disparities become even more striking when individual judge's records are compared. For instance, when responding to mobbing claims, some French appellate judges are over 20 times more (which means others are 20 times less) rigorous than their colleagues. Allegedly. Unfortunately, no statistical evidence can be presented here, as individual statistics are no longer available. The recent French statute - the same which provides for mandatory digital publication of all judicial decisions - makes it a criminal offence to analyze or to compare individual judge's work outcome using big data tools⁴. The author (and the software provider) would therefore risk up to 5 years in prison (and a 300 000 € fine) if they did so...

As shocking as it is in terms of justice transparency⁵, this brings us to the second point. If French state is ready to prosecute for analyzing public servants' activity, it means that for some reasons we are afraid of the truth about judicial decision-making. The question is whether we should be afraid, and what risks are involved.

3. The Dangers of Legal Realism.

What realists had already told us, and what numbers confirm today, is that judicial decision is determined to a very high extent by judge's individual approach. What is at stake here is our understanding of the very nature of judicial review. It seems obvious today, as Richard Posner once put it, that judges act like "occasional legislators"⁶. Judges do not simply apply, they make the law, at least from time to time. When

⁴ Loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice, Article 33 (JORF n°0071 du 24 mars 2019, texte n° 2).

⁵ For a critical comment on the new statute, see: T. PERROUD, *L'anonymisation des décisions de justice est-elle constitutionnelle ? Pour la consécration d'un principe fondamental reconnu par les lois de la République de publicité de la justice*, blog.juspoliticum.com (in French).

⁶ R. POSNER, *How Judges Think*, 2010, Cambridge Mass. - London, p. 78.

choosing between competing solutions, they act as politicians do, as their choice is rarely clearly dictated by the Rule of Law. What big data shows today, is that even when the statute itself is clear, it is applied in a different manner by different courts. The question is: how should we respond to the new evidence of judicial subjectivity?

We may try to reject, or to bury the truth, like French legislators did, but it seems both difficult in a long-term perspective and dubious from the constitutional point of view. Justice is rendered “in the name of the French People”, judges are public servants and, according to Article 15 of the French Bill of Rights, the “Society has the right to request account from public agents”. If we can verify how a given Member of Parliament voted during his or her tenure, we should also be able to verify how a given judge applied the statute voted by the Parliament.

Instead of criminalizing research, we should rather accept and try to deal with the truth about judicial subjectivity.

The risk, however, is a further politicization of our approach to the judiciary. Justice Kavanaugh’s case is only the latest episode in a series of confirmation hearings dominated by political controversy. As controversial as they were at the time, Robert Bork’s rejected nomination, in 1987, as well as Justice’s Clarence Thomas’ confirmation hearings, dominated by sexual harassment accusations, may only announce what is yet to come. Once we acknowledge the intrinsic political nature of judicial decision-making, it seems difficult to separate law from politics. As Critical Legal Studies success in the U.S. shows, this could ultimately lead us to a general rejection of the Rule of Law as a legitimate tool for settling human affairs (not to mention the degree of political controversy we may observe within the American academia). At the same time, the realist argument may be very dangerous when used by some “populist” governments, seeking to undermine the authority and the independence of the judiciary. Some recent attacks on the Rule of Law, in countries like Hungary or Poland⁷, were based on a somewhat realist

⁷ See: W. ZAGORSKI, *Commission/Pologne: la Cour de Luxembourg vient au secours des juges polonais*, blog.juspoliticum.com (in French).

rhetoric, portraying judges as political actors attempting to steal People's sovereignty.

Finally, the ultimate triumph of legal realism could completely shake up our educational model. If legal reasoning relies not on the knowledge of rules, but on the aptitude to assert the situation and to make a proper policy choice, we should wonder if we still need lawyers after all. In a case involving expropriation for public utility, an economist would be better equipped to evaluate all pros and cons of a construction project disputed by the applicant. In a criminal case, a psychologist is better placed to examine mental health of the offender. In social policy matters, statistics could be more helpful than the knowledge of precedent. If law is not about rules, but all about choices, why should we rely on lawyers in the first place?

THE ROLE OF HUMAN JUDGE IN JUDICIAL DECISIONS.
PRELIMINARY REMARKS ON LEGAL INTERPRETATION
IN THE AGE OF ARTIFICIAL INTELLIGENCE

Laura Vagni

TABLE OF CONTENTS: 1. The task of judging and the option of an artificial judge. – 2. Law and interpretation or law as interpretation. – 3. The influence of the human judge on judicial decisions. – 4. The role of the judge: a constant tension between human being and being human.

1. The task of judging and the option of an artificial judge.

We live in a digital age where digital technologies invest most of our daily activities and where Artificial Intelligence (AI) is increasingly integrated in our society¹. The use of algorithms in the judicial process has been questioned since the last years of the past century and is supported by the idea of preventing the mistakes and biases of the judge.

Nowadays, AI introduces the possibility of developing systems of predictive justice and eventually to attributing the task of solving controversies to an artificial judge.

AI is able to collect a huge number of cases and to compare

¹ The meaning of “Artificial Intelligence” and the use of this expression are wide debated in literature. It is outside the scope of the present paper to describe the main thesis on the point. For the first references, I limit myself to referring to the work of S. J. RUSSELL - P. NORVIG, *Artificial Intelligence: A Modern Approach*, 3rd ed., Harlow, 2016, p. 1 ff.; cf. D. HAREL, *Computers Ltd: What They Really Can't Do*, Oxford, 2000, reprinted in 2012, where on p. 194 the Author explains: «The very phrase “Artificial intelligence” [...] seems to be a contradiction in terms. We tend to view intelligence as our non-programmable and hence non-algorithmic characteristic. To many people the very idea of an intelligence machine doesn't sound quite right».

a fact with thousands and thousands of preceding ones with a very narrow margin of error, thus exceeding human skills in managing a great amount of data and comparing facts². Moreover, AI has the capacity to learn from examples and improve its performance without humans explaining how to accomplish the task it is given³. Thus, it is potentially autonomous and unaffected by contextual conditions.

Many experiments show the ability of artificial intelligence to predict the solution of a case better than humans. Predictive software has been used in the United States since the last years of the past century and some European systems⁴ are also experimenting with it.

The recent EU General Data Protection Regulation provides specific guarantees for the protection of personal data, especially when used for profiling and automated decision-making, and protects the right of a person not to be subjected to a completely automated decision process⁵. Art. 22 of the Regulation states: «The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her». The article introduces significant exceptions to the right, such as in the case of an explicit consent by the data subject and the case of a decision «[...] authorized by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights

² D. DOBREV, *The Human Lawyer in the Age of Artificial Intelligence: Doomed for Extinction or in Need of a Survival Manual*, in *J. Int'l Bus. & L.*, 18, 2018, p. 39 ff., at 42; S. MASON, *Artificial Intelligence: Oh Really? And Why Judges and Lawyers Are Central to The Way We Live Now - But They Don't Know It*, in *Computer and Telecommunications Law Review*, 2017, p. 213 ff., at 217.

³ Cf. H. SURDEN, *Machine Learning and Law*, in 89 *Washington Law Review*, 2014, p. 87 ff.

⁴ Cf. L. LARRET-CHAHINE, *La justice prédictive*, in this volume, p. 161 ff.

⁵ The analysis of the EU General Data Protection Regulation is outside the scope of the present paper, for a comment on the topic cf. S. HÄNOLD, *Profiling and Automated Decision-Making: Legal Implications and Shortcomings*, in M. Corrales - M Fenwick - N. Forgó, *Robotics (Eds.), AI and the Future of Law*, Singapore, 2018, p. 123 ff. and references *ivi*; G. RESTA, *Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza*, in *Politica del diritto*, 2019, vol. 2, p. 199 ff., at 204.

and freedoms and legitimate interests [...]». The Regulation, therefore, aims to preserve human dignity and avoid a dehumanised judicial process, but it does not prevent the judge from substantially basing his decision on the solution provided by AI. The issue raises a set of questions, including the influence of a solution given by AI on the judicial process of knowledge and the implicit cognition⁶ generated by the interplay between AI and a human judge. In this context, one main problem arises from the use of AI in creating “risk assessments” to aid judges in making decisions⁷. Indeed, some algorithms used in US courts proved to be “remarkably unreliable” in forecasting events such as recidivism in violent crimes. The COMPAS system (Correctional Offender Management Profiling for Alternative Sanctions), used to predict recidivism in many US lower criminal courts, turned out to be discriminatory on the basis of race. In fact, the COMPAS risk assessment is based upon information gathered from the defendant’s life on the bases of questions and an interview with the defendant covering data concerning the defendant’s criminal life, demographic information and whether anyone in the subject’s family has ever been arrested or is divorced⁸. Questions include also the subject’s high school grades and moral hypotheticals, such as whether the subject agrees or disagrees with some statements. Although the questions on which the risk assessment is based do not show a bias in themselves, the results given by the Algorithm were revealed as gravely biased against black defendants and had high percentage of error.

The probability of bias in the COMPAS risk assessment was admitted by the courts; nevertheless, the use of COMPAS was upheld underlining that it consists in a support for the judge’s decision and not in a substitution of the judge. In the case *State of Wisconsin v. Loomis*⁹, decided in 2016, the Wisconsin Supreme Court was asked to deal with the compatibility of the

⁶ See *infra*, para. 3 and references *ivi*.

⁷ LORD HODGE, *Law and Technological Change*, speech at the British Irish Commercial Bar Association, 9 April 2019, available at <https://www.supremecourt.uk/docs/speech-190404.pdf>.

⁸ Cf. S.K. KATYAL, *Private Accountability in the Age of Artificial Intelligence*, in *UCLA L. Rev.*, 66, 54, 2019, p. 85.

⁹ *State of Wisconsin v. Eric L. Loomis*, 13 July 2016, 881 N.W. 2d 749.

use of COMPAS as a tool for sentencing in the circuit courts. The system was used by the Circuit Court to deny a defendant's motion for post-conviction relief. The decision was not made automatically by the Algorithm: the Court ordered an investigation into the facts to evaluate the risk of recidivism. The investigation included also the consultation of COMPAS risk assessment. The Algorithm established that the defendant presented a high risk of recidivism.

The Circuit Court stated that the algorithm solution was used only to corroborate its finding, and that it would have imposed the same sentence regardless of whether it considered the COMPAS risk scores. Moreover, the Court was aware of the risks related to using COMPAS when sentencing. Indeed, the report on the investigation on risk of recidivism presented to the Court included a description of how the COMPAS risk assessment should be used and cautioned against its misuse.

The defendant appealed the decision, on the basis of the violation of his right to a due process of law, under three different aspects: «[...] it violates a defendant's right to be sentenced based upon accurate information, in part because the proprietary nature of COMPAS prevents him from assessing its accuracy; it violates a defendant's right to an individualised sentence; and it improperly uses gendered assessments in sentencing»¹⁰.

The Wisconsin Supreme Court concluded that the Algorithm can be used when sentencing, but it established methodologies and limitations for the use of COMPAS, in order to bring it into line with the right of a due process.

Firstly, the right of the defendant to be sentenced based upon accurate information is to be protected through an adequate awareness of the courts concerning the risk of COMPAS errors. The courts have to be apprised of some problems, namely the need for the tool to be monitored and re-normed for accuracy, due to the change of population; the risk that the algorithm will give a disproportional classification of a minority offender as having a higher risk of recidivism; the fact that the decision of the algorithm is based on a comparison with a national sample, so it does not limit

¹⁰ *State of Wisconsin v. Eric L. Loomis*, cit., at 757.

the evaluation to cases from the same State; the fact that the proprietary nature of the algorithm may prevent it from disclosing how risk scores are to be determined. This information should encourage a cautionary approach to the solutions proposed by the algorithm.

Secondly, the Court stated that the right of an individualised sentence is protected every time the decision is not the result of an automated process, but the court analyses the results given by COMPAS together with other factors, so as to achieve the best solution for the case.

Finally, on the issue of gender discrimination, the Court recognised that COMPAS risk scores take gender into account, however, the Court stated that the failure to distinguish between men and woman misclassify both genders, achieving more unfair results: «Thus, if the inclusion of gender promotes accuracy, it serves the interests of institutions and defendants, rather than a discriminatory purpose»¹¹.

The cautions and limitations in the use of COMPAS, established by the Wisconsin Supreme Court in the *Loomis* case, seem unable to prevent any algorithm bias in risk assessment. In fact, bias can creep in at many stages of the AI learning process: not only at the stage of the collection of data (it may happen that data collected are unrepresentative of reality or reflect existing prejudice) but also in the preparation stage, when the attributes that the algorithm has to consider are selected, and during the processing¹². The method of data processing is difficult to control, as it acquires autonomy even from the programmers who determine the processing rules. Again, these rules are quite difficult to access and understand for the courts. Finally, the influence of the algorithms on the cognitive process of the judge is still underestimated.

The problem of AI bias has recently been approached by identifying the ethical by design as a priority and premise of the application of AI in judicial systems.

In the first European Ethical Charter on The Use of Artificial Intelligence in Judicial Systems and Their Environment, the need for an ethical by design approach is

¹¹ *State of Wisconsin v. Eric L. Loomis*, cit., at 766.

¹² Cf. S.K. KATYAL, *op. cit.*, p. 89.

clearly highlighted¹³. The first principle of the Charter is focused on the respect for fundamental rights to «[...] ensure that the design and implementation of artificial intelligence tools and services are compatible with fundamental rights». The comment states:

«When artificial intelligence tools are used to resolve a dispute or as a tool to assist in judicial decision-making or to give guidance to the public, it is essential to ensure that they do not undermine the guarantees of the right of access to the judge and the right to a fair trial [...] They should also be used with due respect for the principles of the rule of law and judges' independence in their decision-making process. Preference should therefore be given to ethical-by-design or human-rights-by-design approaches»¹⁴.

These approaches are to be encouraged, as the protection of human rights heavily depends on the ethical safeguards in the interplay between human and machine. There is no certainty, however, that they will succeed in assuring the principle of fair trial. One problem may derive from the difficulty in containing the influences exercised by AI on the judge. From this perspective, the right to access the algorithm and be informed of how risk scores are determined, although important, does not always assure the independence of judge. Another problem relates to the possibility of developing an adequate ethical-by-design approach for judicial decision. One may wonder if the ethical principles that the judge (and the AI who supports the judge) needs to make a decision can be fixed and projected in advance or if they need to be interpreted according to the context where the judge gives the judgment. The question is if principles and values can be abstracted or, conversely, if they result from a constant mediation between law and reality. Law is a contextual science and fairness in law depends also on specific

¹³ European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment, Strasbourg, 3 - 4 December 2018, available at <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>.

¹⁴ *Ibidem*.

contextual elements: judicial interpretation cannot develop outside and irrespective of specific circumstances, isolated from a precise time and cultural dimension¹⁵. Consequently, the ability of AI to dissociate itself from contexts, which is the feature that assures its independence and neutrality, at the same time risks being one cause of unfairness in an AI decision. The issue brings to mind the question: «How and to what extent does judicial interpretation relate to the evaluation of the contextual elements of the case?».

Ultimately, the question is: «How and to what extent does judicial interpretation presume a human judge?». The answer to this complex problem concerns the idea of justice we want to embrace and the values and the virtues in law that we want to preserve and pursue. It invests the influence of the judge's humanity in decision-making and the role of the judge in decision making.

2. Law and interpretation or law as interpretation.

The inquiry into the role of the human being in judicial decisions requires an investigation of the interplay between law and interpretation and the mutual influence between law and who interprets law. The issue recalls an age old theme that has been tackled in different ways in all legal experiences, in the course of history, and can be expressed in a concise question: «Does the law exist without interpretation?». Traditionally, the problem is approached differently by common law and civil law systems, the former assuming the idea of law as practice, the latter considering the law as a science.

The English experience is one of the best examples among the European legal systems of an approach toward law as a practice, which has a set of implications on the relationship between law and judicial interpretation. On the one hand, English common law, as case law, has been developed over centuries as judicial interpretation; in modern times, the recognition of the doctrine of binding precedent did not change the nature of case law as a discursive law. Judicial precedents do not consist of legal definitions: they are binding rules but they conserve their nature

¹⁵ D. DOBREV, *op. cit.*, p. 64.

of illustrative and explanatory principles enunciated by judges for a particular case. Thus, «The reason and spirit of cases make law, not the letter of particular precedents»¹⁶. On the other hand, statutory law, that today covers the majority of English law, acquires its effectiveness through statutory interpretation. Since the early common law, the law incorporated in cases has always been the principal reference for statutory interpretation. In *The Institutes of the Laws of England*, Lord Coke explained: «[...] the surest construction of a statute is by the rule and reason of common law»¹⁷.

The English rules of statutory interpretation are tools elaborated by judges for judges, in order to communicate the meaning of the text of a statute and to convert the general statement of a statute into a concrete rule to be applied in the case under judgement. In the course of history, the approach of judges towards statutory interpretation has also led to the attempt to use equity as a canon of interpretation and give the statute the “internal sense of it”, instead of that emerging from the words of the text. This idea, which has always been tempered by the profound respect for the principle of Parliamentary sovereignty, can also be traced in some theories on statutory interpretation up to the eighteenth century. So, Blackstone wrote:

«[...] the most universal and effectual way of discovering the true meaning of a law, when the words are dubious, is by considering the reason and spirit of it; or the cause which moved the legislator to enact it. For when this reason ceases, the law itself ought likewise to cease with it [...]. From this method of interpreting laws, by the reason of them, arises what we call *equity*; which is thus defined by Grotius, “the correction of that, wherein the law (by reason of its universality) is deficient”»¹⁸.

¹⁶ It is the famous opinion of Lord Mansfield in the case *Fischer v. Prince*, 3 Burr 1363, that is still valid today, cf. C.K. ALLEN, *Law in the Making*, 7th ed., Oxford, 1964, p. 216 ff.

¹⁷ SIR. E. COKE, *The First Part of the Institutes on the Laws of England; or, a Commentary upon Littleton*, 18th ed., London, 1794, III, chap. 8, para. 272.

¹⁸ W. BLACKSTONE, *Commentaries of the Law of England*, 1765-69, London, vol. 1, p. 61.

In the nineteenth century, the English judges' attitude towards statutory interpretation was restrained in accordance with the legalistic tendency, and the adherence to the literal meaning of the text was considered a safeguard of the supremacy of Parliament. The relationship between law and interpretation established in that period, however, conserved the fundamental idea that legislation is not enacted in a vacuum, but it presumes common law. The true meaning of legislation is discovered through its application to concrete cases, where legislation becomes experience, filtered by judicial reasoning¹⁹.

On the Continent, conversely, the influence of legal positivism led to the recognition of a supremacy of legislation over law: law is legislation incorporated in the text and interpretation is a bare application of the text. The main reference is to the idea of interpretation expressed by Montesquieu, according to whom judges, like automatons, «[...] are only the mouths that pronounce the words of the law, inanimate beings, who can moderate neither its force nor its rigor»²⁰.

The myth of law without interpretation had already been dashed against reality by the end of the nineteenth century. In the French system, the law professor Geny originated the free scientific research movement in jurisprudence; he criticised the traditional approach toward the interpretation of written law and recognised the creativity of judicial activity. The judge has to shape the decision on the basis of a free search of the rule, that is «[...] outside the reach of any positive authority»²¹. At the same time judicial research needs to be “[...] objective, because it can be solidly based upon objective elements which systematic-scientific jurisprudence alone can reveal”²². Along these lines, Geny underlined that there is no

¹⁹ Z. BANKOWSKI - D. N. MACCORMICK, *Statutory Interpretation in the United Kingdom*, D. N. MacCormick - R.S. Summers (Eds.), *Interpreting Statutes. A Comparative Study*, Aldershot, 1991, p. 360.

²⁰ C. L. DE SECONDAT MONTESQUIEU, *The Spirit of the Laws*, translated and edited by A. M. Cohler - B.C. Miller - H.S. Stone, Cambridge, 1989, p. 163.

²¹ F. GÉNY, *Méthode d'Interprétation et Sources en Droit Privé Positif*, 2d. ed., English translation by the Louisiana State Law Institute, Paris, 1954, p. 355, para. 156.

²² *Idem*.

system of interpretation that can flatter itself for having completely eliminated the personal evaluation of the interpreter; thus, the conscience and the intuition of the judge are involved in the process of interpretation: «[...] the principles revealed by conscience and recognised by human reason through intuition are the first necessary directive for the free search of an interpreter of positive law»²³.

In the first decades of the XX century, the attack on formalistic legal interpretation found decisive support in the movement of legal realism, developed among American scholars and then widespread in both common law and civil law countries.

Legal realism emphasised the role of humanity in judicial decisions and went far as admitting that they have an irrational component²⁴. Holmes, who is considered the initiator of the movement, stated that: «The life of law has not been logic; it has been experience»²⁵, pointing out a close relationship between law and interpretation and between law and experience. Firstly, the relationship between law and interpretation is to be analysed according to a logic of inclusiveness. Secondly, there is an intimate relationship between law and experience: in the judicial process law becomes experience.

Law as experience involves the humanity of the judge, not only his rationality.

Every judicial decision implies a double process of interpretation: on the one hand, the judge is asked to read the facts of the case, selected and produced during the process by the parties; on the other hand, the judge has to interpret law, in order to provide justice in the concrete case, to find the right solution for the case.

The facts of the case impact and influence the cognitive process of the judge and the law itself that the judge is asked to interpret. The Judge, like a scientist, is involved in the study of phenomena. However the judge, unlike the observers of other physical phenomena, deals with the behaviour of

²³ *Ibidem*, p. 365, para. 160.

²⁴ V. TUMONIA, *Legal Realism and Judicial Decision-Making*, in *Jurisprudence*, 2012, 19(4), p. 1361 ff.

²⁵ O. W. HOLMES, *The Common Law*, New York, [1881] 1991, p. 1.

human beings²⁶. Thus, the task of judging implies human relationships and this has an influence on the process of study.

The idea of a judge that applies the law to the facts, according to a linear process of interpretation, starting from the reading of the appropriate rule and ending with the judicial decision, when the rule is applied to the facts, seems to be unrealistic and even misleading. Conversely, the judicial decision follows a circular process: the impact of the judge on the facts leads to a first pre-comprehension of the question submitted to him. The relationship between the judge and parties, in the case in question, forges the judge's first interpretation. The initial comprehension is then checked through a process of reinterpretation and revision of the first idea. The cognitive process of the judge, therefore, is a round trip²⁷. Along these lines, law is interpretation and not a premise of interpretation. Any phase of this movement stimulates the human nature of the judge in the interplay with the parties of the process and the facts of the case, as well as in analysing the rules and checking his first intuition.

3. The influence of the human judge on judicial decisions.

The question concerning the humanity of the judge and its involvement in the judicial decision has long been investigated in legal literature, with opinions differing as to the possibility and, if necessary, the opportunity of correcting its effects on the judgement. The issue was at the centre of the legal realism movement, mentioned above. In this scenario, the studies of Roscoe Pound on the influence of non-legal elements in the decision-making process assume major importance. The Author states that there is a discretionary element in judicial justice:

«Two antagonistic ideas, the technical and the discretionary, may be seen at work throughout the administration of

²⁶ W.W. COOK, *The Logical and Legal Basis of the Conflict of Laws*, in *Yale L. J.*, 33, 1924, p. 457, at 475.

²⁷ L. MOCCIA, *Riflessioni sparse (e qualche involontario aforisma) su interpretazione e diritto*, in L. Moccia, *Comparazione giuridica e prospettive di studio del diritto*, Lavis, 2016, p. 47.

justice. These might well be called the legal and the non-legal element in judicial administration. [...] in no legal system, however minute and detailed its body of rules, is justice administered wholly by rule, without any recourse to the will of the judge and his personal sense of what should be done to achieve justice in the cause before him. Both elements are to be found in all administration of justice»²⁸.

In a famous paper published in 1931 Jerome Frank, commenting on Pound's theory, questioned if judges are human and criticised the possibility that they might distance themselves from their human passions and character²⁹. Human nature is intimately mingled with judicial technique and the immense importance of personal elements in court justice cannot be ignored. Whereas according to other scholars, the judge has the ability to distinguish between law and what law is not, departing from social, economic, political and physiological influences. All these pressures are conducive to judicial discretion, but rules and discretion in a judicial process are separable:

«The way in which we can predict how a judge will decide an issue on which a rule of law exists is after all different from the way in which we can anticipate from observed past uniformities how a muscle will react when brought into contact with an electric current. In the case of the muscle, prediction does not have to assume that the muscle will react with conscious intent to conform to a rule formulating its proper behaviour under the circumstances. That is, however, precisely what a judge can and must in most cases be expected to do»³⁰.

Jerome Frank criticised the attempt to ask the judge to think

²⁸ R. POUND, *Justice According to Law*, in *Columbia Law Review*, 13, 1913, p. 696.

²⁹ J. FRANK, *Are Judges Human? The Effect on Legal Thinking of the Assumption That Judges Behave Like Human Beings*, part I, in *U. Pa. L. Rev.*, 80, (1931-1932), p. 17 ff.

³⁰ J. DICKINSON, *Legal Rules: Their Function in the Process of Decision*, in *U. Pa. L. Rev.*, 1930-1931, 79, p. 833, at 839.

in an artificial way and stated that: «He [the judge] will be ashamed of the way his mind works humanly despite his efforts. He will waste precious hours attempting to think inhumanly [...]»³¹.

The judge is first of all a human being who looks at the world through the lens of his own experience, education, culture and vision of life. His humanity filters into the judicial decision and influences the development of the judicial process. The personality of the judge, his preconceptions about the issues that he is asked to judge and about the parties of the process as well as his emotions influence the meaning of the facts subjected to his interpretation.

The debate developed by the legal realism movement in the first decades of the last century is extremely relevant in our post-modern age, where the role of the judge as an actor of the development of law, instead of a bare applicator of legislation, clearly emerges. The multiple and multilevel sources of law, on the one hand, and the fast-changing social reality, on the other, put the task of judging outside the comfortable binaries and taxonomies of legal positivism and invest the judge with the complex task of confronting legal diversity and even with contrasting views of legal phenomena. The system challenges the technical and legal skills of the judge, but also his character and personality.

An Italian Constitutional court judge recently affirmed that the syllogistic process does not describe the judicial interpretation in our post-modern age³². Conversely, in the decision the judge has to search for the solution and this research constitutes a discovery that implies a process that is inconsistent with syllogism. In this process the rationality of the judge and his/her logical capacity are involved, but also his intuition, perception and comprehension are all relevant at an axiological level.

Good character has always been considered a fundamental feature of judges in the common law tradition, where the

³¹ JEROME FRANK, *Are Judges Human?* (...), cit., p. 24.

³² P. GROSSI, *La invenzione del diritto: a proposito della funzione dei giudici*, in *Riv. Trim. Dir. Proc. Civ.*, 2017, p. 831; L. MOCCIA, *Riflessioni sparse (e qualche involontario aforisma) su interpretazione e diritto*, cit., p. 43 ff., at 50.

development of case law was due predominantly to judges of deep knowledge and strong personality. Even recently Lady Hale, in a speech about *Moral Courage in the Law*, explained that an effective judiciary needs well qualified and learned judges³³. However, a judge with these essential characteristics is not necessarily a good judge: he or she also needs to show moral courage. Moral courage has different meanings, including «[...] courage to recognise and stand up to our own prejudices, preconceptions and predispositions – and the first step is recognising that we may have them»³⁴.

The contemporary legal literature commonly admits that the judge's temperament has a certain degree of influence on the solution of the case. So the task of judging «[...] is likely to be influenced by temperament, emotion, experience, personal background and ideology (influenced in turn by temperament and experience), as well as by an 'objective' understanding of what should be the 'best' legislative policy to adopt in order to resolve the issue in the case»³⁵.

In this scenario, a new contribution derives from neuroscience research applied to law and judicial decision. Neuroscientists revealed that the personality of the judge and his/her emotional character influence the cognitive process of judicial decision, at a level that the judge's rationality cannot control. Some studies show that judges can fall into implicit biases, into cognitive traps, and how their decisions are deeply influenced by affective sensations³⁶.

The myth of the blind judge recalls the idea of pure rationality in decision-making, whereas the development of cognitive science demonstrates that the decision is partly dependent on the judge's human nature, and it is impossible to

³³ LADY HALE, *Moral Courage in the Law*, The Worcester Lecture 2019, Worcester Cathedral, 21 February 2019, available at <https://www.supremecourt.uk/docs/speech-190221.pdf>; LORD CLATKE, *Selecting Judges: Merit, Moral Courage, Judgment and Diversity*, in *High Ct. Q. Rev.*, 2009, 5(2), p. 49, speaking about the need for judges to have good character: «Moral courage rather than moral cowardice is needed for good character to be satisfied».

³⁴ LADY HALE, *op. ult. cit.*, p. 8.

³⁵ R.A. POSNER, *How Judges Think*, Harvard, 2010, p. 173.

³⁶ Among the wide literature, cf. A. BETHOZ, *Emotion and Reason: the Cognitive Neuroscience of Decision Making*, Oxford, 2006.

isolate intellectual functions in a human being, without considering their interplay with emotions and sensations. Judges are subject to implicit, or hidden, cognition when they make a choice. Among the implicit cognition that impacts the judicial decision, neuroscience particularly underlines bias, emotion and empathy³⁷.

The “emotional” judge has been seen as a peril for justice and due process³⁸. Conversely, scientific and technological progress may appear to be an antidote against the risk of discretion in the administration of justice. Along these lines, the use of algorithms in judicial decisions, as a partial or total substitute for the human judge, suggests a model of neutral and objective justice, more conformant with the myth of the blind judge, who decides on the basis of pure rationality, a possibility definitively dispelled by neuroscience research.

The option of an “artificial” judge raises not only the question concerning the contribution of technology to judicial justice, but also the need for and utility of humanity in judicial decisions.

4. The role of the judge: a constant tension between human being and being human.

The judicial process is the greatest example of the mutual relationship between law and interpretation: law is interpretation and, during the process, law becomes interpretive judicial practice. The relationship between law and interpretation is not linear, but circular and the entire judicial decision is dominated by a continuous shift between the humanity of the judge and the role of judging: the impact of facts on the humanity of the judge and his role of researching the law for the case. Therefore, the judicial interpretation implies a constant tension between human being and being human³⁹. It is inevitable that

³⁷ A. S. BRADLEY, *The Disruptive Neuroscience of Judicial Choice*, in *UC Irvin Law Review*, 9(1), 2018, p. 1 ff., at 7.

³⁸ Cf. for example the study by A. FORZA - G. MENEGON - R. RUMINATI, *Il Giudice emotivo. La decisione tra ragione ed emozione*, Bologna, 2017, p. 207.

³⁹ The expression echoes the study by A. J. HESCHEL, *Who Is Man?*, Stanford, 1965, p. 27.

the judge approaches the decision on the basis of an initial view and intuition, however, comparing his initial view with the legal and policy arguments, he/she could depart from his/her initial view. In other words, the judge needs the ability to self-correct: this is an important feature of the role of the judge, as it helps to achieve the right conclusion when, for example, there are more possible results for a case⁴⁰. The first interpretation of the judge is influenced by his/her own personality, but the humanity of the judge, is closely connected with the dignity of the task of judging.

The tension between human being and being human requires the recovery of the cultural dimension of judging and the role of the judge as a cultural mediator, neglected by a technical and scientific idea of law. The digital age makes it even more urgent to recover the humanity of the judge, and with it the human dignity of the role. To pursue this aim we need firstly to clear the field of an ambiguity that afflicts our time: the task of solving cases and the task of judging cases are not the same, even if their results usually overlap. The scientific and legal knowledge, the capacity to collect, classify and compare data, are important skills for solving the case, but they need to be accompanied by the ability of the judge to interpret law. This is a human ability, as it needs awareness of the contextual dimension of law; in other words, it needs humanity: a free will that impacts with the concrete facts of the case together with the responsibility to seek justice for that case.

The judge is a product of his/her time and interprets it with the awareness of a man/woman who moves in a logical system of probable knowledge. In this context, the judge makes full use of his/her experience to understand the current times and ways of our society, which are continuously transforming. Furthermore, the judge should be immersed in society, instead of alienated from it or unaffected by its ideas and forces. The role of the judge requires a constant fight between what is scientifically certain and what is fair and aspiring towards

⁴⁰ LORD NEUBERGER, *Judge Not, That Ye Be Not Judged: Judging Judicial Decision-Making, AF Mann lectures, British Institute of International and Comparative Law*, 29 January 2015, available at <https://www.supremecourt.uk/docs/speech-150129.pdf>.

justice⁴¹. During a lecture at the British Institute in 2015, Lord Neuberger affirmed:

«[...] we must approach the task of judging in a manner which embraces, rather than eschews, our humanity. We should do so more openly and more honestly».

From this perspective, the debate about the use of automatons and artificial intelligence in the decision-making process should be approached accepting the humanity of the judge instead of denying it.

Technological progress provides a great opportunity for challenging the role of the human judge and recovering the reasons for its dignity. In our time, as in the past, this process needs education and human virtues: that is the comprehension of technological progress but also the courage to be part of it, without rejecting it.

Technological progress and the development of artificial intelligence ask us to revolutionise the education of judges to help them understand a changing world.

The common law countries and, especially the US, are a driving force in reforming legal education programmes so as to include in the curricula of the law schools law and technological teaching, work on legal practice technologies, legal design projects, and innovation incubators⁴². Some scholars, however, affirm that the approach towards the inclusion of technological science into legal curricula are sometimes characterised by a sort of “anxious legal studies”⁴³: the urgency for the legal control and regulation of

⁴¹ L. MOCCIA, *op. ult. cit.*, p. 53.

⁴² L. FLORIDI (Eds.), *The Onlife Manifesto. Being Human in a Hyperconnected Era*, 2015, p. 9: «Throughout our collective endeavour, a question kept coming back to the front stage: “What does it mean to be human in a hyperconnected era?” This foundational question cannot receive a single definitive answer, but addressing it has proven useful for approaching the challenges of our times»; L. MOCCIA, *Education to Intercultural Citizenship: a European Perspective to Global Citizenship*, in *La Cittadinanza Europea*, 2014, p. 161, at 177: «Notwithstanding sophisticated devices that in few seconds make people reachable everywhere, we still need a humanised common set of values and habits of mind upon which to base peaceful and prosperous coexistence in the world».

⁴³ Cf. J. WEBB, *Information Technology & the Future of Legal Education*:

technology. This approach leads to an interpretation of the relationship between law and technology along binary lines: melding technology and law as far as we can, pushing the technological courses, or “weaponising” the law to resist the technological revolution. They affirm, on the contrary, that the entanglement of humans and technology needs to be seriously investigated, evaluating how to best deal with the responsibility deriving from this relationship. In addition, technological learning should not be marginalised in legal education, but «[...] it must be understood and problematized pervasively across the curriculum»⁴⁴.

Technological progress challenges the entire vision of legal education because it challenges the idea of what it is to be human and, consequently the role of humanity in law. Therefore, the need to take seriously the interplay between AI and law in legal education implies learning to be human. We need to discover and exercise those human virtues in order to live together in a digital age.

Paradoxically, one main challenge for judicial interpretation in the age of AI is to tackle the question: «What does it mean to be human in a judicial decision?».

a Provocation, in *Griffith Journal of Law and Human Dignity*, Special Issue, 2019, p. 1 ff.

⁴⁴ At a practical level, legal education would encourage the introduction in the curricula of subjects to develop skills of design thinking, critical thinking and creative-problem solving, design oriented teaching and research, ethical and governance approach, cf. *Ibidem*, p. 21 ss.

Il giurista si trova sempre più impegnato a confrontarsi con gli scenari inediti che si schiudono con la massiccia utilizzazione degli algoritmi e, in generale, delle nuove tecnologie. Uno degli aspetti più rilevanti della diffusione degli impieghi dell'intelligenza artificiale è che con essa sembra incrinarsi, se non proprio spezzarsi, uno dei cardini del pensiero giuridico moderno, costituito dalla possibilità di imputare una condotta ad un soggetto in base alla autonomia del processo decisionale che ne è alla base. La scelta di concentrarsi sul tema della decisione nasce dal tentativo di individuare una chiave di lettura che consenta di attraversare trasversalmente la complessità e la vastità dei profili connessi allo sviluppo delle nuove tecnologie nel mondo del diritto, cogliendone uno dei nodi più delicati. I contributi raccolti nel presente Volume ruotano attorno a questa area tematica, muovendo da un approccio interdisciplinare, che si apre al dialogo con la scienza informatica e con la filosofia, per concentrarsi poi su alcuni aspetti nella prospettiva storico-giuridica, privatistica, di diritto dei trasporti, di diritto amministrativo e di diritto comparato, con particolare riguardo anche alle nuove frontiere della cd. giustizia predittiva.

ERMANNO CALZOLAIO è professore ordinario di Diritto Privato Comparato nel Dipartimento di Giurisprudenza dell'Università di Macerata. È coordinatore del progetto "Dipartimento di Eccellenza" dal titolo "Diritto e Innovazione: Europa e Cina di fronte alle sfide della globalizzazione". Svolge attività didattica e di ricerca presso numerosi atenei stranieri, tra cui l'Università di Orléans (Francia), la Shanghai University of Political Sciences and Law (Cina), la Beijing Normal University (Cina), la Murdoch Law School (Australia), l'Università di Lund (Svezia), la Jagiellonian University (Cracovia). È membro dell'International Academy of Comparative Law e componente del Comitato Direttivo dell'Associazione Italiana di Diritto Comparato. I suoi interessi di ricerca riguardano il ruolo del precedente nella comparazione civil law-common law e nel diritto europeo, nonché il diritto della proprietà, dei contratti, della responsabilità civile e delle successioni nella prospettiva comparatistica.



€ 25,00 I.V.A. INCLUSA