

INDICE

	<i>pag.</i>
<i>Introduzione</i>	1

PARTE I

UNA TECNOLOGIA *UNPREDICTABLE BY DESIGN*: POTENZIALITÀ E RISCHI DELL'INTELLIGENZA ARTIFICIALE

CAPITOLO I

L'INTELLIGENZA ARTIFICIALE: CARATTERISTICHE E SVILUPPO DI UNA NUOVA TECNOLOGIA

1. L'intelligenza artificiale, un concetto in cerca di definizione, tra innovazione tecnologica e suggestioni fantascientifiche	17
2. Dall'i.a. simbolica al connessionismo in tre fasi: logica, conoscenza, apprendimento	22
2.1. Intelligenza come applicazione deduttiva di principi logici: l'i.a. simbolica degli anni '50-'60	23
2.2. Intelligenza come applicazione deduttiva di conoscenze specialistiche: gli <i>expert systems</i> degli anni '80	25
2.3. Intelligenza come apprendimento induttivo e probabilistico: l'attuale successo del <i>machine learning</i>	26
3. Limitazione della nostra indagine ai sistemi di <i>machine learning</i> : il problema del controllo (rinvio)	31
4. Anche i sistemi di i.a. sbagliano: errori, vulnerabilità e limiti degli algoritmi di apprendimento automatico	32
4.1. Il problema della "scorciatoia" e l' <i>adversarial machine learning</i>	33
4.2. La discriminazione algoritmica	36

CAPITOLO II
**L'INTELLIGENZA ARTIFICIALE
 E IL PROBLEMA DEL CONTROLLO**

1.	Il problema del controllo in una prospettiva socio-tecnologica	43
1.1.	L'autonomia come capacità di agire in situazioni di incertezza	49
1.2.	L'automazione come capacità di svolgere attività in assenza di intervento umano	53
1.3.	L'opacità tecnologica	56
1.3.1.	L'opacità del modello algoritmico: inesplicabilità del <i>decision making</i> artificiale	57
1.3.2.	L'opacità del "sistema socio-tecnologico" nel suo complesso	61
2.	Una precisazione sul problema del controllo: imprevedibilità del singolo <i>output vs</i> monitoraggio delle <i>performance</i> complessive. Inaccettabilità teorica delle tesi che concepiscono l'i.a. come "magia" o "seconda natura"	65
3.	A cavallo tra fenomeno e giudizio normativo: incapacità dei sistemi di i.a. di agire liberamente	73
3.1.	(<i>segue</i>) Incapacità di agire in maniera consapevole e di recepire il significato dei comandi giuridici	78

CAPITOLO III
**INTELLIGENZA ARTIFICIALE
 E FATTI ILLECITI: UNA CASISTICA**

1.	Una panoramica sulla casistica del "danno da intelligenza artificiale"	83
2.	Gli <i>embodied systems</i> : i sistemi dotati di <i>hardware</i>	86
2.1.	Veicoli a guida autonoma	86
2.2.	Sistemi di guida autonoma degli aerei	97
2.3.	Sistemi d'arma autonoma letali	100
2.4.	<i>Healthcare robots</i>	105
3.	I <i>bots</i> : i sistemi informatici in grado di effettuare una serie di compiti automatizzati	109
3.1.	<i>Social bots</i>	110

	<i>pag.</i>
3.2. <i>Bots</i> di acquisto <i>online</i>	113
3.3. Sistemi di <i>trading</i> algoritmico	114
4. Gli <i>intelligent decision support systems</i> : il supporto decisionale ad individui e imprese	118
4.1. Sistemi di supporto alla diagnostica e al trattamento medico	119
4.2. Sistemi di supporto decisionale nei contesti bellici	124
5. L'intelligenza artificiale generativa: la creazione algoritmica di contenuti originali	126
5.1. <i>Chatbot</i> basati su modelli di linguaggio di grandi dimensioni (<i>large language models</i>)	126
5.2. Piattaforme per la generazione di <i>deepfakes</i>	130

PARTE II

LA REGOLAZIONE DELL'INTELLIGENZA ARTIFICIALE: UNO SGUARDO INTERDISCIPLINARE

CAPITOLO IV

IL REGOLAMENTO EUROPEO SULL'I.A. (AI ACT) E I PRIMI TENTATIVI DI REGOLAZIONE IN ITALIA E ALL'ESTERO

1. La regolazione dell'intelligenza artificiale. Una panoramica sulle strategie a livello internazionale: Europa, Stati Uniti, Cina, Regno Unito	137
2. Il regolamento europeo sull'intelligenza artificiale (c.d. AI Act)	149
2.1. Il campo di applicazione del regolamento: la definizione di "sistema di i.a."	151
2.2. La classificazione dei sistemi di i.a. sulla base del rischio	156
2.3. (<i>segue</i>) I requisiti per lo sviluppo e la messa in commercio dei sistemi di i.a. ad alto rischio	164
2.3.1. Il principio di trasparenza	165
2.3.2. Il principio della sorveglianza umana	171
2.3.3. Gestione dei dati, robustezza, accuratezza e cibersecurity	174

	<i>pag.</i>
2.4. La gestione del rischio da intelligenza artificiale e gli obblighi di monitoraggio	176
2.5. L'armonizzazione europea delle norme tecniche: profili generali e applicazione al settore dell'intelligenza artificiale	180
2.6. <i>Governance</i> e vigilanza sull'applicazione dell'AI Act: il ruolo delle autorità di vigilanza sul mercato	185
3. La regolazione dell'intelligenza artificiale in Italia: le (poche) disposizioni già in vigore e il d.d.l. n. S. 1146 del 2024 contenente «disposizioni e delega al Governo in materia di intelligenza artificiale»	193
3.1. Profili penalistici del d.d.l. n. S. 1146 del 2024	201
4. Riepilogo: le strategie di regolazione dell'imprevedibilità algoritmica	213

CAPITOLO V

LA RESPONSABILITÀ CIVILE EXTRACONTRATTUALE PER I DANNI “DA INTELLIGENZA ARTIFICIALE”

1. La responsabilità civile per i danni da intelligenza artificiale: l'allocatione del costo economico del danno come questione di politica legislativa	222
1.1. Dal dogma della colpa alla responsabilità oggettiva per la creazione del rischio: alcuni cenni storici	225
1.2. Ritorno al futuro: un declino del modello oggettivistico? Regolazione dell'intelligenza artificiale e <i>policy</i> europea	231
1.3. Un dilemma noto: adattabilità della normativa vigente o necessità di introdurre una disciplina <i>ad hoc</i> ?	234
2. La responsabilità del produttore di sistemi di i.a. La nuova direttiva 2024/2853/UE sul danno da prodotto: un adattamento dei requisiti tradizionali, con un inedito cuore processuale	238
2.1. L'estensione del campo di applicazione della disciplina: le nuove nozioni di “prodotto” e “produttore”	242
2.2. La difettosità del prodotto e l'esimente per il rischio di sviluppo: responsabilità oggettiva o cripto-colpa? Una perimetrazione del concetto di “difetto”	246
2.2.1. La classificazione tradizionale applicata ai “prodotti intelligenti”: difetti di fabbricazione, progettazione, informazione	251

	<i>pag.</i>
2.2.2. Verso una nuova nozione di <i>design defect</i> in relazione ai sistemi di i.a.? La direttiva 2024/2853/UE, tra nuovi parametri di valutazione ed esimenti rimaste inalterate	258
2.2.3. Le proposte avanzate dalla dottrina: l'agente modello umano, l'agente modello algoritmico, la sostanziale eliminazione del requisito del difetto	264
2.2.4. I rapporti tra <i>product safety</i> e <i>product liability</i> : il prodotto conforme ai requisiti di sicurezza è necessariamente "non difettoso"? Sull'esistenza di un presunto "rischio consentito" in ambito civilistico	267
2.2.4.1. La tesi positiva: la c.d. <i>preemption doctrine</i> . Cenni a una teoria incompatibile con l'approccio europeo alla sicurezza dei prodotti	270
2.2.4.2. La tesi negativa. Le ragioni della necessaria separazione tra <i>product liability</i> e <i>product safety</i>	273
2.3. L'accertamento del nesso di causalità. Alcune note sul rapporto tra causalità generale e causalità individuale nell'ambito della responsabilità per danno da prodotto	276
2.4. Il cuore processuale della direttiva 2024/2853/UE: le presunzioni di difettosità e di causalità	287
2.5. La responsabilità del produttore di sistemi di i.a. per l'esercizio di attività pericolose (art. 2050 c.c.)	291
3. La responsabilità oggettiva in capo all'utilizzatore del sistema di i.a.: sull'applicabilità dei regimi speciali previsti dal codice civile	293
3.1. Responsabilità da cosa in custodia (art. 2051 c.c.) e responsabilità del custode dell'animale o di chi se ne serve (art. 2052 c.c.)	295
3.2. Responsabilità da circolazione di veicoli (art. 2054 c.c.)	298
3.3. Responsabilità di padroni e committenti (art. 2049 c.c.)	300
4. La personalità elettronica: una proposta europea presto abbandonata	302
5. <i>Manufacturer enterprise liability</i> e assicurazione obbligatoria	306
6. Riepilogo. Intelligenza artificiale e responsabilità extracontrattuale: indicazioni e spunti per il prosieguo dell'indagine	309

PARTE III
**INTELLIGENZA ARTIFICIALE
 E RESPONSABILITÀ PENALE: UNA SISTEMATICA**

CAPITOLO VI
DELIMITAZIONE DELL'INDAGINE PENALISTICA

1. Soggetto o oggetto del diritto? L'intelligenza artificiale nell'ordinamento penale, a metà via tra <i>res cogitans</i> e <i>res extensa</i> . L'inaccettabilità delle teorie sulla responsabilità penale diretta dei sistemi di i.a. (rinvio)	321
2. Il nostro contributo allo studio del tema indagato e al dibattito esistente: una parte generale della responsabilità penale "da intelligenza artificiale"	325
3. La tassonomia utilizzata: la separazione tra reati colposi e reati dolosi come principale criterio di classificazione	327
4. Una panoramica esemplificativa sulle fattispecie legali configurabili	331
4.1. Reati colposi e intelligenza artificiale: i reati contro la vita e l'integrità fisica e i reati contro l'incolumità pubblica	332
4.1.1. Reati colposi di danno	332
4.1.2. Reati colposi di pericolo	334
4.2. Reati dolosi e intelligenza artificiale: un ampio orizzonte di nuove "opportunità" criminose	335
4.2.1. Reati dolosi di danno	336
4.2.2. Reati dolosi di pericolo	339
5. La struttura dei prossimi capitoli. Qualche ulteriore precisazione sulle scelte operate	342
5.1. La responsabilità amministrativa da reato degli enti: le ragioni di un'esclusione "eccellente" (rinvio)	344
6. Rischio <i>noto</i> o <i>ignoto</i> ? Il rapporto tra intelligenza artificiale e principio di precauzione	348

SEZIONE I
LA RESPONSABILITÀ COLPOSA

CAPITOLO VII
IL DANNO DA PRODOTTO “INTELLIGENTE”
E LA RESPONSABILITÀ COLPOSA
DEL PRODUTTORE

1. La responsabilità colposa del produttore per l'evento lesivo “da intelligenza artificiale”. Premessa	358
2. La struttura del fatto tipico nella responsabilità penale per danno da prodotto: la “fungibilità”, nell'applicazione giurisprudenziale, tra commissione e omissione	361
3. L'accertamento del nesso di causalità. La “flessibilizzazione” emersa nella giurisprudenza sulla responsabilità per danno da prodotto tradizionale	370
3.1. Nesso di causalità e intelligenza artificiale. Analogie e distinzioni rispetto alla casistica sul danno da prodotto tradizionale	379
3.1.1. La c.d. <i>black box</i> algoritmica: problemi di imputazione dell'evento lesivo nell'ambito delle organizzazioni complesse (rinvio)	382
3.1.2. L'opacità “complessiva” del sistema di i.a.: problemi di multifattorialità	387
3.2. L'autonomia dei sistemi di i.a. come possibile fattore interruttivo del nesso di causalità?	392
4. La responsabilità individuale nel contesto di una criminalità intrinsecamente plurisoggettiva: il c.d. <i>many hands problem</i>	398
4.1. Il <i>many “technical” hands problem</i> : l'imputazione dell'evento lesivo in capo a ricercatori e tecnici	400
4.2. Il <i>many “decisional” hands problem</i> : l'imputazione dell'evento lesivo in capo agli appartenenti al management della società. Cenni critici alla nuova veste “assorbente” della posizione di garanzia nelle organizzazioni complesse	405
5. La colpa del produttore di sistemi di i.a. per il verificarsi dell'evento lesivo algoritmico	415

	<i>pag.</i>
5.1. Una panoramica sulle regole cautelari scritte in materia di i.a.: dall'assenza di regolazione alla consueta "iperregolazione"	418
5.1.1. Alcune tendenze delle cautele scritte in contesti di incertezza scientifica: la funzione "precauzionale" e la funzione "burocratico-organizzativa"	425
5.2. Il rapporto tra regole cautelari scritte e regole cautelari non scritte: quale spazio per il rischio consentito?	432
5.2.1. Individuazione delle aree di rischio consentito e limiti all'operatività dell'istituto. Il c.d. "fallimento" delle regole cautelari scritte	442
5.2.2. Le regole cautelari non scritte: la colpa generica come inevitabile "rinforzo" di quella specifica?	448
5.3. La prevedibilità dell'evento: può considerarsi <i>prevedibile</i> l' <i>imprevedibilità algoritmica</i> ?	451
5.3.1. (<i>segue</i>) Principio di precauzione e giudizio di prevedibilità dell'evento: una "relazione pericolosa" incompatibile con il <i>risk-based approach</i> adottato dall'AI Act	460
5.4. Il nesso colpa-evento di fronte alla <i>black box</i> algoritmica	466
6. Intelligenza artificiale e imputazione colposa degli eventi lesivi in capo al produttore: riepilogo	469

CAPITOLO VIII

LA RESPONSABILITÀ DELL'UTILIZZATORE-SUPERVISORE DEL SISTEMA DI I.A.

1. La responsabilità del supervisore di un sistema di i.a. e la c.d. " <i>moral crumple zone</i> "	475
2. Il principio della sorveglianza umana e l'introduzione di obblighi di supervisione	480
3. La struttura del fatto tipico. Gli obblighi di supervisione: regole cautelari o norme istitutive di posizioni di garanzia?	484
4. L'accertamento del nesso di causalità (rinvio)	488
5. La colpa. Una panoramica sulle regole cautelari prospettabili	490
6. Le cautele che regolano l'interazione tra utilizzatore-supervisore e sistema di i.a. dotato di automazione. Il caso della circolazione di auto guida autonoma	496

	<i>pag.</i>
6.1. Un possibile affidamento sulla “diligenza algoritmica”?	503
6.2. Obbligo di controllo e potere di controllo: il rischio di una frattura strutturale	508
6.3. Le prospettive <i>de iure condendo</i> . L’introduzione di regole cautelari scritte rigide: il perimetro del rischio consentito	516
7. Le cautele che regolano l’interazione tra utilizzatore-supervisore e sistema di i.a. privo di automazione	520
7.1. Il caso dei sistemi di supporto all’attività clinica e diagnostica dei medici	523
7.2. Equiparazione degli <i>output</i> algoritmici alle linee guida?	530
8. Intelligenza artificiale e imputazione colposa degli eventi lesivi in capo al supervisore: riepilogo	537

SEZIONE II LA RESPONSABILITÀ DOLOSA

CAPITOLO IX L’UTILIZZO ILLECITO DEL SISTEMA DI I.A.: LA RESPONSABILITÀ DOLOSA DELL’UTILIZZATORE

1. L’intelligenza artificiale nei reati dolosi: strumento <i>vs</i> esecutore materiale del reato	543
2. L’utilizzo illecito di un “prodotto” generato dall’intelligenza artificiale: da ChatGPT ai <i>deepfakes</i> . Un problema di regolazione	548
3. L’utilizzo illecito dell’attività svolta da un sistema di i.a. Il difficile inquadramento del “fatto” algoritmico nella condotta tipica dell’agente	554
3.1. L’oggetto del dolo e l’imprevedibilità algoritmica	558
3.1.1. La divergenza tra voluto e realizzato in relazione al decorso causale: <i>l’aberratio causae</i>	559
3.1.2. La divergenza tra voluto e realizzato in relazione all’evento: <i>aberratio ictus</i> e <i>aberratio delicti</i>	561
3.1.3. La programmazione “generica” dell’algoritmo: l’ir-rilevanza del <i>dolus generalis</i>	568
4. Il privilegio dell’utilizzatore di attività algoritmica illecita: riepilogo e prospettive <i>de iure condendo</i> (rinvio)	577

CAPITOLO X

**LA RESPONSABILITÀ DOLOSA DEL PRODUTTORE PER GLI
OUTPUT LESIVI DEI SISTEMI DI I.A.**

- | | | |
|------|--|-----|
| 1. | La responsabilità dolosa del produttore per gli <i>output</i> lesivi derivanti da un sistema di i.a. (generativa, in particolare) | 583 |
| 1.1. | Il produttore come concorrente nel reato commesso dall'utilizzatore | 587 |
| 1.2. | Il produttore come autore del reato doloso | 590 |
| 2. | L'inapplicabilità del meccanismo di <i>notice and action</i> previsto dal Digital Services Act | 592 |
| 3. | Intelligenza artificiale e generazione di informazioni false. Spunti dalla giurisprudenza civile: il caso canadese relativo al <i>chatbot</i> di una compagnia aerea e la giurisprudenza in materia di <i>Autocomplete</i> di Google | 595 |
| 3.1. | (segue) Spunti dall' <i>enforcement</i> del GDPR in relazione ai sistemi di i.a. generativa. Il diritto di rettifica e di cancellazione dei dati personali | 601 |
| 4. | Riepilogo. La prevenzione degli <i>output</i> illeciti dei sistemi di i.a. generativa: il controllo svolto dalle autorità amministrative | 605 |

PARTE IV

**CONCLUSIONI E PROSPETTIVE
SUL RAPPORTO TRA DIRITTO PENALE
E INTELLIGENZA ARTIFICIALE**

CAPITOLO XI

**CRITICA ALLE TEORIE DELLA RESPONSABILITÀ
PENALE DIRETTA DEI SISTEMI DI I.A.**

- | | | |
|----|---|-----|
| 1. | Il dibattito sulla responsabilità penale diretta dei sistemi di i.a. come <i>stress test</i> per le categorie penalistiche | 613 |
| 2. | La "tesi utilitaristico-funzionalista" di Simmler e Markwalder: la soggettività di diritto penale come costruito esclusivamente socio-normativo | 617 |

	<i>pag.</i>
2.1. Critica: incompatibilità delle premesse utilitaristico-funzionalistiche con la concezione personalistica della responsabilità penale	620
3. La “tesi biofisica” di Hallevy: la configurabilità di <i>actus reus</i> e <i>mens rea</i> nell’agire algoritmico	624
3.1. Critica: il sistema di i.a. come soggetto incapace di essere destinatario dei comandi penali	627
3.2. (<i>segue</i>) Il sistema di i.a. come soggetto incapace di colpevolezza	631
4. « <i>A body to kick, but still no soul to damn</i> »: critica delle pene “robotiche”	634
5. Il paragone fuorviante con la responsabilità da reato degli enti	639
5.1. Differenza “topografica” rispetto al diritto penale delle persone fisiche: affiancamento <i>vs</i> sovrapposizione	640
5.2. Differenza di <i>ratio</i> : deterrenza mediata <i>vs</i> deterrenza immediata	642
5.3. Differenza fenomenica: substrato socio-normativo <i>vs</i> substrato naturalistico	644
6. Gli obiettivi politico-criminali dietro le proposte di responsabilità penale diretta dei sistemi di i.a.: un incentivo all’azzardo morale?	646

CAPITOLO XII
**RUOLO E LIMITI DELL’INTERVENTO
 PENALISTICO NEL SETTORE
 DELL’INTELLIGENZA ARTIFICIALE**

1. Intelligenza artificiale ed esigenze politico-criminali: riepilogo. “Ordinare il molteplice”: la necessità di un sistema di tutela integrato, tra diritto amministrativo, diritto civile e diritto penale	650
2. La definizione del rischio consentito: il contesto geopolitico e il ruolo degli “enti di normalizzazione” nell’emissione degli <i>standard</i> tecnici	657
2.1. (<i>segue</i>) Le tecniche di standardizzazione. Dal <i>Code is law</i> alla <i>Governance by numbers</i> : un approccio quantitativo al rischio consentito	662

	<i>pag.</i>
3. La gestione del rischio da intelligenza artificiale: la <i>governance</i> europea. Le <i>regulatory sandboxes</i> e la valutazione delle autorità di vigilanza sulla sicurezza del sistema di i.a. Ancora sull'approccio quantitativo al rischio consentito	670
3.1. (<i>segue</i>) L'autoregolazione e l'auto-organizzazione degli enti	676
4. Il <i>côté</i> compensativo: la tutela di diritto civile extracontrattuale	679
5. Il <i>côté</i> preventivo-repressivo: la tutela di diritto amministrativo. Le sanzioni per l'omessa comunicazione di informazioni rilevanti sul rischio	682
6. (<i>segue</i>) La tutela penalistica	688
6.1. La responsabilità penale del produttore: crisi del diritto penale d'evento e tecniche di anticipazione della tutela penale	690
6.1.1. Prospettive <i>de iure condito</i> : le fattispecie previste dall'art. 112 cod. cons. per la messa in commercio di prodotti pericolosi per la vita e l'incolumità individuale	697
6.1.2. Prospettive <i>de iure condendo</i> . La messa in commercio di sistemi di i.a. pericolosi per la vita e l'incolumità individuale	703
6.1.3. (<i>segue</i>) La mancata predisposizione di adeguati meccanismi di sorveglianza umana	708
6.1.4. (<i>segue</i>) Il coordinamento tra attività dell'autorità di vigilanza e tutela penale: il modello ingiunzionale e l'omessa comunicazione delle lesioni e dei decessi derivanti dai sistemi di i.a.	711
6.2. La responsabilità colposa dell'utilizzatore-supervisore: la possibile rilevanza settoriale del principio di affidamento "robotico"	718
6.2.1. La limitazione della colpa alla sola forma grave: criticità e potenzialità	722
6.3. La responsabilità dolosa dell'utilizzatore: l'anticipazione della tutela penale in alcuni specifici settori	731
7. La responsabilità da reato degli enti: potenzialità e limiti del d.lgs. 231/2001 nel settore dell'intelligenza artificiale	737

	<i>pag.</i>
8. Le prospettive aperte dal d.d.l. n. S. 1146 del 2024 contenente «disposizioni e delega al Governo in materia di intelligenza artificiale»: spunti per un'integrazione del testo di legge e per un'eventuale attuazione delle deleghe	747
9. Chiusa. Il diritto penale e la politica criminale al tempo della trasformazione algoritmica della società	757
<i>Bibliografia</i>	761