BIG DATA

Innovazione, giustizia, protezione dati

IL CONTESTO

L’impatto del digitale sulla giustizia riguarda essenzialmente la digitalizzazione dell’attività giudiziaria e l’inclusione, nell’oggetto della giurisdizione, delle norme di fonte principalmente europea, per la regolazione dell’ecosistema digitale.

La digitalizzazione della giustizia è un obiettivo che la Comunicazione del 2020 della Commissione europea assegna tanto alle istituzioni europee per promuovere l’efficacia della cooperazione giudiziaria (v. Reg. (2023) 2844, quanto ai singoli Stati membri, sostenendone il percorso di innovazione per migliorare l’efficacia dei sistemi giudiziari e per migliorare e facilitare l’accesso alla giustizia da parte dei cittadini e delle imprese

DIGITALIZZAZIONE

- La riforma Cartabia con i suoi regolamenti attuativi ha valorizzato in misura significativa la telematizzazione del processo, mentre il PNRR ha promosso la costituzione di basi informative importanti
- Peraltro, la digitalizzazione non tocca solo il profilo organizzativo e strumentale ma anche quello più strettamente processuale, investigativo e probatorio. Soprattutto su questo terreno, il ricorso alla tecnologia e alle sue potenzialità crescenti lascia intravedere l'esigenza di una più puntuale regolazione, come emerso in relazione **alla *data retention* e ai criptofonini**: temi sui quali la giurisprudenza, europea e interna, ha dovuto svolgere un'azione per certi versi di supplenza, per altri di monito al legislatore.

LA REGOLAZIONE DEL DIGITALE

Norme unionali recenti (Digital Markets Act, Digital Services Act, Data Governance Act, Data Act) hanno tentato di disciplinare l'ecosistema digitale per un verso responsabilizzando (DMA e DSA) le piattaforme in funzione della *cybersafety* e della concorrenza e, per altro verso, promuovendo (DGA e DA) la creazione di uno spazio europeo dei dati, da condividere a fini tanto solidaristici quanto economico-produttivi.

La responsabilità, civile e penale derivante dalla violazione di tali norme è ovviamente rimessa all'a.g. degli Stati membri

PROTEZIONE DATI

- Per entrambi gli aspetti, le intersezioni con la protezione dati sono importanti, anche in ragione dell' applicabilità, dal 2018, della fonte normativa europea non solo all'ambito della cooperazione giudiziaria (in cui significativamente anche il Reg. 2844 la richiama) ma, più incisivamente, a quello dell' attività giudiziaria interna, in maniera diretta per le giurisdizioni diverse da quella penale e, per quest'ultima, con la mediazione della normativa di recepimento della direttiva (UE) 2016 (680) (d.lgs. 51 del 2018).

- La duplicità dei plessi normativi applicabili rende evidente, già di per sé sola, la complessità del rapporto tra giustizia, digitale e privacy, che inizieremo ad affrontare anzitutto sotto il profilo delle recenti politiche d'innovazione.

BIG DATA

I “big data” (ovvero grandi volumi di dati, provenienti da fonti eterogenee, assistiti da caratteristiche di veridicità e suscettibili di analisi di particolare velocità) di derivazione giudiziaria o, comunque, funzionali alla giurisdizione presentano, infatti, caratteristiche tali da esigere cautele peculiari e garanzie rafforzate nella loro utilizzazione, per la tutela:

- dei soggetti interessati (parti processuali, terzi…) e
- degli stessi interessi pubblicistici sottesi (si pensi, per tutti, al segreto investigativo o all’autonomia e indipendenza della magistratura).

GIUSTIZIA E DIGITALIZZAZIONE

- La riforma Cartabia (d.lgs. 149 e 150 del 2022) in entrambi i settori della giustizia ordinaria e il PNRR valorizzano l'uso delle risorse digitali in ambito giurisdizionale, che tra gli sviluppi più recenti annovera progetti quali la piena realizzazione del processo telematico, in ambito sia civile che penale, la costituzione di una serie di sistemi telematici funzionali alla giurisdizione quali, ad esempio, gli albi dei ctu o, più recentemente, le infrastrutture digitali per le intercettazioni, recentemente centralizzate per garantire maggiore efficienza e, al contempo, sicurezza, nella gestione di flussi informativi così delicati.
- Gli aspetti più delicati, dal punto di vista della protezione dati, comuni a questi progetti, pur nella loro diversità, possono ricondursi a due macroaree: **la sicurezza e la riservatezza.**

SICUREZZA

Comune ai vari progetti d'innovazione è l'esigenza di sicurezza dei flussi informativi (conformemente al principio di cui all'art. 5, p.1, lett. f) GDPR nonché all'art. 3, c.1, lett.f) d.lgs. 51 del 2018), prevedendo misure tecniche e organizzative realmente adeguate al grado di rischio connesso al trattamento (artt. 32 GDPR; 25 d.lgs. 51 del 2018), come:

- accessi selettivi ai dati da parte dei soli soggetti legittimati in ragione della funzione svolta;
- tracciabilità delle operazioni effettuate per ricostruirne poi le dinamiche in caso di accessi illeciti;
- procedure affidabili per la risoluzione di eventi critici come i *data breach* o i "*disaster*" per i quali approntare sistemi di "*recovery*" efficienti;
- presidi essenziali quali, ad esempio, la crittografia.

SICUREZZA E COOPERAZIONE GIUDIZIARIA

Il Regolamento sulla digitalizzazione nell'ambito della cooperazione giudiziaria (n. 2844) qualifica, significativamente, la sicurezza dei sistemi informatici utilizzati (oltre alla loro interoperabilità) quale requisito necessario per gli scambi informativi previsti (cfr., in particolare, art.3, pp.1 e 4).

La sicurezza, integrità e attendibilità dei flussi documentali e l'identificazione dei partecipanti alla comunicazione sono parametri valutativi dell'adeguatezza dei canali tecnologici utilizzati (C19).

PROPORZIONALITA’

Anche la proporzionalità della raccolta informativa (che corrisponde ai principi di minimizzazione, limitazione della finalità e della conservazione di cui agli artt. 5 GDPR e 3 d.lgs. 51 del 2018) contribuisce alla garanzia di maggiore sicurezza del trattamento, nella misura in cui riduce la superficie esposta a potenziali attacchi.

Ciò implica, in particolare, la raccolta dei soli dati necessari (e non eccedenti rispetto) al fine sotteso al trattamento e la conservazione (in chiaro) per il solo tempo necessario (con cancellazione o anonimizzazione successiva)

LINEE DI RESPONSABILITA'

La definizione dell'architettura del trattamento può rivelarsi particolarmente complessa per quanto attiene alla demarcazione delle attribuzioni tra Ministero e uffici giudiziari, nella misura in cui la **centralizzazione dei sistemi informativi attrae la competenza verso il primo** a fronte, tuttavia, della **titolarità dei secondi rispetto ai trattamenti correlati all'ambito processuale.**

Così, ad esempio, mentre dei trattamenti realizzati mediante vari sistemi informativi centrali è generalmente titolare il Ministero della giustizia, del trattamento realizzato, anche su dati raccolti mediante questi portali, ma nell'ambito del singolo procedimento, è titolare l'ufficio giudiziario (cfr. circolare del Ministero della giustizia prot. n. 21611 del 27 giugno 2018). Ancora, mentre per gli albi distrettuali dei ctu è titolare l'ufficio giudiziario competente, per l'elenco nazionale la titolarità è ascritta al Ministero.

SOGGETTI DEL TRATTAMENTO

- L’articolazione del trattamento dei dati in segmenti diversi si riflette in un’architettura giocata su forme diverse, che prevedono secondo i casi titolarità autonome, contitolarità o relazioni tra titolare e responsabile. E se questa complessità è, ormai, molto frequente soprattutto nelle pubbliche amministrazioni, i cui procedimenti implicano trattamenti di dati personali dalle filiere articolate con segmenti talora autonomi, nel settore giudiziario essa riflette anche i particolari valori e assetti costituzionali sottesi.
- Ad esempio, rispetto alle **infrastrutture digitali per le intercettazioni**, l’attribuzione della titolarità agli uffici giudiziari e il divieto, normativamente sancito, di accessibilità dei dati in chiaro al Ministero riflettono l’esigenza di garanzia non solo del segreto investigativo ma, anche, dell’autonoma direzione delle indagini da parte della magistratura requirente

LINEE DI RESPONSABILITÀ NELLA COOPERAZIONE

Articolazione analoga (in termini di rapporto tra centralizzazione e decentramento) rispetto a quelle su descritte si ha anche in ambito di cooperazione giudiziaria europea.

L’art. 14 del Regolamento 2844 imputa, ad esempio, alla Commissione la titolarità dei trattamenti realizzati mediante il punto di accesso elettronico europeo e alle autorità nazionali competenti la titolarità dei trattamenti dei dati inviati o ricevuti tramite il sistema informatico decentrato.

PUBBLICITÀ E RISERVATEZZA

Determinante nei processi di digitalizzazione è anche il bilanciamento tra **pubblicità degli atti procedimentali** (amplificata esponenzialmente dal mezzo telematico) e **garanzia della riservatezza** delle parti e dei terzi coinvolti.

Un profilo peculiare è emerso a seguito dell'estensione, con il d.lgs. 150 del 2022 del ricorso, nel processo penale alla **riproduzione audiovisiva e fonografica come modalità generale di documentazione**, che è destinata ad affiancare il verbale per gli atti del procedimento (art. 134 c.p.p.), quale modalità preferenziale di documentazione dell' interrogatorio di garanzia dell'indagato *in vinculis* (art. 141-*bis* c.p.p.), quale forma di documentazione dell'assunzione dibattimentale dei mezzi di prova (art. 510, c.2-*bi*s c.p.p.).

PUBBLICITÀ E RISERVATEZZA NELLA COOPERAZIONE

Un’analoga esigenza di bilanciamento tra pubblicità degli atti e riservatezza si riscontra negli artt. 5, p. 3 e 6, p.7, del Regolamento 2844, rispetto alla registrazione delle udienze in videoconferenza, con, in particolare per l’ambito penale:

- l’ulteriore, necessaria garanzia di riservatezza dei **colloqui tra difensore e assistito**, richiamata dall’art. 6, p.5 e
- la **tutela del superiore interesse del minore** di cui all’art. 6, p. 6 .

DIGITALIZZAZIONE DEL PATRIMONIO INFORMATIVO

Un ulteriore aspetto delle politiche di innovazione riguarda la digitalizzazione del patrimonio informativo del sistema-Giustizia, per la quale il PNRR ha previsto, tra l'altro:

- la progettualità di un *data-lake* inteso come punto di accesso unico "*a tutti i dati grezzi prodotti dal sistema giudiziario*", ordinario e amministrativo, nonché
- la creazione di una banca dati gratuita e accessibile di tutte le decisioni (per ora realizzata rispetto alle pronunce civili di merito).

BDP

- Il database delle decisioni civili di merito, pubblicamente accessibile è stato realizzato con pronunce risalenti a non più di otto anni fa, previa pseudonimizzazione dei dati identificativi delle parti, con i nomi dei magistrati in chiaro ed escludendo i provvedimenti ad anonimizzazione obbligatoria ex art. 52, c.5, dlgs. 196 del 2003 (in materia di minori, famiglia, stato della persona).
- Si coniugano così esigenze di pubblicità (e accessibilità) del patrimonio informativo giudiziario e tutela della riservatezza. Il regime ordinario di pubblicità delle sentenze (art. 51-52 d.lgs. 196 del 2003), fondato sulla regola della pubblicità e sull'oscuramento come eccezione, sarebbe infatti potuto apparire non del tutto adeguato al grado di rischio connesso alla raccolta, centralizzata in unico data base e dunque con una possibilità di incroci di dati e inferenze ben ulteriori rispetto a quelle proprie della pubblicazione sui siti istituzionali del singolo ufficio giudiziario.

IL RISCHIO DI REIDENTIFICAZIONE

- Tra i rischi del trattamento, da valutare ai sensi degli artt. 32 Gdpr e 25 dlgs 51/18 ai fini dell'adozione di misure di contrasto, vi è quello della reidentificazione degli interessati, da impedire in particolare nei casi in cui l'anonimato sia dovuto, come per le ipotesi di cui all'art. 52, c. 5, dlgs 196 del 2003, significativamente non censite nella banca dati di merito.
- Tale esigenza è tanto più rilevante in relazione ai "dati grezzi" che dovrebbero rifluire nel data-lake, non necessariamente assistiti, dunque, dal regime di accessibilità proprio delle decisioni giurisdizionali ai sensi dell'art. 51, c.2, d.lgs. 196 del 2003.
- Si potrebbe anche sfruttare – come suggerisce la stessa Comunicazione 2020 della Commissione - le potenzialità dell'i.a. per anonimizzare i provvedimenti da includere nel data-lake, utilizzando algoritmi allenati a riconoscere gli elementi da oscurare

LIMITI DELL'A.I.

E' difficile, però, pensare a una delega totale alla macchina dell'anonimizzazione di dati solo indirettamente identificativi (indirizzi, numeri telefonici, numeri di conti bancari, codici fiscali, ecc,) con particolare riguardo a quelli ad oscuramento obbligatorio ex art. 52, c.5, rispetto ai quali l'apprezzamento sulla capacità di disvelare l'identità del soggetto è particolarmente complesso.

L'analisi del rischio di reidentificazione è, significativamente, uno degli adempimenti previsti, in Francia, come preliminari alla pubblicazione delle decisioni giudiziarie nella c.d. Loi Numerique, volta a coniugare certezza dell'interpretazione giuridica con la sua necessaria vitalità.

E come ricordava la stessa CEPEJ nella Carta etica del 2018, "*non è stato ancora ideato*" (ed è probabile non lo sia stato neppure ora) "*un meccanismo automatizzato di anonimizzazione ex post pienamente efficace, che impedisca qualsiasi rischio di identificazione e riidentificazione"*.

IN PROSPETTIVA…

Sul punto **sarebbe auspicabile una disciplina più puntuale**, che identifichi la tipologia di dati raccolti, gli algoritmi applicati, il carattere aperto o meno del formato in cui i dati sono resi disponibili (anche alla luce delle perplessità avanzate proprio dal CEPEJ sul punto), le garanzie da adottare, le modalità di esercizio, da parte degli interessati, dei diritti loro riconosciuti dal GDPR (da assicurare laddove non possa assicurarsi un'anonimizzazione almeno sufficiente ai sensi del C 26) e le misure volte a contenere- sulla scorta della stessa legge francese- il rischio di profilazione dei magistrati, pur nel rispetto del principio di pubblicità del processo di cui all'art. 6, par.1, CEDU.

Andrebbe poi prevista la supervisione umana sull' ipotesi di anonimizzazione proposta dall'algoritmo, al fine di colmare eventuali carenze dovute ai limiti di comprensione della macchina e correggerne ulteriormente l'apprendimento automatico.