L’applicabilità della definizione di dato personale ai big data

Determinare l’ambito di applicazione della disciplina di cui al Codice della privacy ed alla direttiva 95/46/CE – che resterà in vigore fino al 25 maggio 2018 quando sarà sostituito dal Regolamento 679/2016 – è un compito complesso quanto essenziale al fine di rendere più efficiente l’attività di trattamento delle informazioni.

Assoggettare un dato alla disciplina prevista dal Codice senza che ve ne sia bisogno porterebbe l’azienda a compiere scelte erronee circa l’opportunità o meno del trattamento dovute alla maggiore onerosità dello stesso. Di contro non fornire adeguata tutela a informazioni personali può portare a sanzioni pecuniarie anche di notevole entità come indicato dal titolo III, capo I del Codice.

L’art. 5 alla rubrica “oggetto ed ambito di applicazione” prevede che “il Codice disciplina il trattamento dei dati personali”. Solo le informazioni che si possono definire tali, infatti, godono di tutte le garanzie previste dalla normativa italiana ed europea in materia di tutela della privacy.

L’ampiezza della definizione della direttiva, da cui prende spunto l’art. 4, comma 1, lett. b) del Codice, garantisce la possibilità agli interpreti di leggere la norma al fine di adeguarla alle novità tecnologiche che si susseguono nel tempo. **Il 16 settembre 2014**, infatti, **il gruppo di lavoro ex articolo 29 ha affermato che “*The EU legal framework for data protection is applicable to the processing of personal data in big data operations*”.**

Da un’analisi approfondita delle norme definitorie, anche alla luce del parere 4/2007 sulla definizione di dato personale del gruppo di lavoro europeo, è possibile scorgere un punto critico molto rilevante ai fini pratici e riscontrabile nel concetto di “identificabilità” dei dati personali.

Ogni informazione può essere definita identificabile nel momento in cui sia possibile rilevare, direttamente o indirettamente, un collegamento pressoché univoco tra questa ed il soggetto cui si riferisce attraverso l’utilizzo dei mezzi che possono essere “ragionevolmente utilizzati”. La ragionevolezza dell’utilizzo dei mezzi discende da un criterio di

utilità puramente economico che esula dal mero tecnicismo. Non basta, dunque, che l’identificazione sia possibile ma è necessario che l’utilizzo abbia un peso non spropositato secondo una logica costi-benefici.

Se le previsioni sulla diffusione dei big data saranno rispettate è ragionevole pensare che in ambito lavorativo questa nuova modalità di analisi dei dati potrà essere utilizzata per determinare un collegamento univoco tra informazioni e soggetto a cui si riferiscono. Inoltre sarebbe difficile anche quantificare l’effettivo vantaggio derivante dallo sfruttamento dei dati che, grazie a questa nuova modalità di analisi, è certamente aumentato in modo esponenziale soprattutto in funzione dell’ingente quantità di dati che è possibile processare. Dato quindi un incremento dei benefici ed una diminuzione dei costi derivanti dal trattamento di dati personali attraverso i big data, **è facile pensare che le nuove tecnologie che permettono questi nuovi processi rientreranno presto – se non l’hanno già fatto – tra i “mezzi che possono essere ragionevolmente utilizzati”.**

Ad una prima analisi parrebbe che le problematiche derivanti dall’uso di tali mezzi siano riferibili solamente a grandi imprese che hanno infrastrutture che permettono di raccogliere elevate quantità di dati. Tuttavia è possibile che tali informazioni siano trasferite ad imprese anche di dimensioni più modeste che, quindi, si trovano a trattare dati non acquisiti direttamente. **Il trasferimento delle informazioni diventa rilevante soprattutto se queste sono anonimizzate poiché in questo caso l’azienda ricevente non è soggetta alla disciplina posta a protezione dei dati personali**. Per questo il parere 5/2014 sulle tecniche di anonimizzazione del *working party* pone l’accento in diversi punti sulla necessità di effettuare un controllo periodico sulla effettiva impossibilità di effettuare un processo inverso a quello di anonimizzazione: il pericolo nascente dal trattamento di dati che potrebbero diventare personali e che sfuggono alla disciplina del d.lgs. 196/2003 è evidente.

Da questo punto di vista il problema principale è dato dall’identificazione c.d. indiretta che proprio coi big data ha acquisito un ruolo centrale in termini di trattamento di un’informazione alla stregua di dato personale o meno. Viene definita identificazione indiretta l’attività volta alla ricerca di un nesso tra alcune informazioni e un soggetto – o un gruppo di soggetti – ben definiti attraverso l’utilizzo di dati ulteriori rispetto a quelli oggetto dell’analisi che

possono provenire dalle fonti più disparate (social network, open data, ecc.).

Data la semplicità con cui è possibile accedere ad informazioni che facilitano l’identificazione indiretta è lecito chiedersi se sia ancora possibile parlare di dati non personali o meno. La risposta non è facile anche in vista di un bilanciamento con l’interesse economico e sociale connesso allo sfruttamento di informazioni che ad oggi sono soggette a regole più permissive e che possono avere anche un’utilità pubblica come avviene, ad esempio, per i dati aggregati che vengono utilizzati nella ricerca scientifica o in ambito statistico. Questi, infatti, possono essere resi pubblici solo se non identificabili nemmeno indirettamente ex art. 8, comma 2, allegato A, d.lgs 196/2003. Ritenere tutte le informazioni assoggettabili alla normativa standard prevista per i dati personali renderebbe questo tipo di operazioni eccessivamente gravose soprattutto per ciò che concerne l’ottenimento di un espresso consenso scritto.

Una delle soluzioni che si potrebbero prospettare sarebbe quella di rendere più certi e determinati i limiti per la ricorrenza del c.d. “interesse legittimo” ex art. art. 24, lett. g) del Codice che, pur aprendo una via economicamente più vantaggiosa, rimane ancora una strada difficilmente percorribile soprattutto per via dei rischi cui si espongono i *data controller* nel trattare dati personali senza una vera e propria autorizzazione scritta.

Gaetano Machì

Studente di Giurisprudenza

Università degli Studi di Milano

@Gae95

Scarica il pdf