

Procedure di e-discovery e tutela dei dati personali: una questione di metodo

Sara Gobbato

Presidente del Tribunale dell'Unione europea, ord. 29 ottobre 2020, T-451/20 R, Facebook Ireland Ltd. c. Commissione europea¹

È sospesa l'esecuzione della decisione della Commissione europea di richiesta di informazioni in materia antitrust ai sensi del regolamento (CE) n. 1/2003, nella parte in cui obbliga l'impresa interessata a fornire documenti contenenti dati personali sensibili di natura medica. La sospensione ha effetto fintanto che l'impresa non attui una specifica procedura che comporta la trasmissione su supporto elettronico separato dei documenti contenenti i dati personali sensibili, da collocarsi in una virtual data room accessibile ad un numero ristretto di funzionari della Commissione e ad un ugual numero di avvocati della stessa impresa. I funzionari dovranno selezionare i documenti in questione ritenuti rilevanti, dando la possibilità agli avvocati dell'impresa di esprimere osservazioni. In caso di disaccordo sulla rilevanza di un documento ai fini del procedimento antitrust, per la composizione della controversia l'impresa potrà ricorrere al Direttore per l'Informazione, la Comunicazione ed i Media presso la DG Comp.

Sommario

1. L'ordinanza del Presidente del Tribunale - 2. Il quadro normativo UE. - 3. Metodologie di e-discovery e *predictive coding*: uno sguardo all'esperienza USA e UK.

Keywords

dati personali – antitrust - e-discovery - *predictive coding* - Commissione europea

1. L'ordinanza del Presidente del Tribunale

Nell'economia digitale le imprese devono gestire non solo grandi quantità di dati ma anche altrettanto ampie moli di documenti, che possono formare oggetto di richieste di informazioni da parte delle pubbliche autorità. In materia antitrust, il regolamento (CE) n. 1/2003² attribuisce alla Commissione europea poteri di indagine che compren-

¹ Tutti i siti web citati nella presente nota sono stati visitati da ultimo il 16 febbraio 2021.

² Regolamento (CE) n. 1/2003 del Consiglio, del 16 dicembre 2002, concernente l'applicazione delle regole di concorrenza di cui agli articoli 81 e 82 del trattato, in *GU L 1* del 4.1.2003, 1–25. L'art. 18, par. 3, stabilisce in particolare che: «Quando richiede alle imprese o associazioni di imprese di comunicare

dono, oltre alle ispezioni, anche le richieste di informazioni rilevanti per l'accertamento di violazioni degli artt. 101 e 102 TFUE da parte delle imprese.

Codificati all'epoca dell'analogico, oggi tali poteri di indagine vengono esercitati nel nuovo contesto digitale dominato dalle Electronically Stored Information (ESI), con l'insorgere di due difficoltà speculari: la Commissione deve individuare criteri di ricerca delle informazioni rilevanti che assicurino la necessità e proporzionalità della richiesta rispetto alle obiettive finalità dell'indagine; le imprese, tenute ad adempiere alle richieste a pena di sanzioni, devono definire procedure interne che limitino i documenti prodotti in risposta all'autorità, evitando la diffusione sia di dati coperti da riservatezza sia di informazioni comunque eccessive rispetto all'oggetto dell'indagine.

Il bilanciamento fra i contrapposti interessi di Commissione e imprese è l'oggetto delle valutazioni espresse dall'ordinanza presidenziale del Tribunale dell'Unione europea del 29 ottobre 2020 in commento. Nel caso di specie, nell'ambito di un procedimento antitrust la Commissione europea aveva ordinato a Facebook - ai sensi dell'art. 18, par. 3, del regolamento (CE) n. 1/2003³ - di fornire documenti interni da selezionarsi sulla base di alcuni criteri di ordine temporale e parole chiave di significato piuttosto ampio⁴, a pena di una sanzione amministrativa pari a €8 milioni al giorno in caso di inottemperanza.

Rispetto alla richiesta di informazioni così formulata, Facebook aveva individuato ben 729.417 documenti. Di questi, 645.459 erano stati forniti alla Commissione. Per i rimanenti 83.958 documenti, tuttavia, l'impresa aveva deciso di richiedere in via cautelare la sospensione della decisione della Commissione nella parte cui essa riguardava documenti ritenuti irrilevanti e non necessari rispetto all'indagine antitrust, contenenti da un lato informazioni riservate sull'attività d'impresa e dall'altro dati personali relativi in particolare alle condizioni di salute dei dipendenti, inclusi nella corrispondenza con i medici curanti nonché nei referti acquisiti dalla società nell'ambito dell'attività di gestione delle risorse umane⁵.

Sotto il profilo del *fumus boni juris* quanto alla conformità della richiesta di informazioni rispetto all'art. 18, par. 3, del regolamento (CE) n. 1/2003, nella propria ordinanza il Presidente del Tribunale riconosce innanzitutto che, in applicazione delle parole chiave individuate nella richiesta di informazioni, Facebook è tenuta ad interrogare i propri server fornendo una grande mole di documenti, dei quali sarà la Commissione a poter stabilire la rilevanza solo in una fase successiva del procedimento. La richiesta di informazioni non prevede né misure idonee a limitare la produzione di documenti irrilevanti, né garanzie a tutela delle persone interessate dalla diffusione dei relativi dati, potendo dunque determinare *prima facie* una violazione dell'art. 18, par. 3, del Regola-

informazioni mediante decisione, la Commissione indica le basi giuridiche e lo scopo della domanda, precisa le informazioni richieste e stabilisce un termine entro il quale esse devono essere fornite. Indica altresì le sanzioni previste dall'articolo 23 e indica o commina le sanzioni di cui all'articolo 24. Fa menzione inoltre del diritto di presentare ricorso dinanzi alla Corte di giustizia delle Comunità europee avverso la decisione».

³ Decisione C(2020) 3011 final del 4 maggio 2020 (caso AT.40628 – Facebook Data-related practices).

⁴ Le parole consistevano in espressioni come “*big question*”, “*for free*”, “*shut down*” and “*not good for us*”. Vd. § 39 dell'ordinanza in commento.

⁵ Vd. ordinanza, § 28.

mento (CE) n. 1/2003 nella misura in cui comporti l'acquisizione di informazioni non necessarie ai fini dello specifico procedimento antitrust⁶.

Quanto all'incidenza della richiesta di informazioni sui diritti fondamentali delle persone interessate, il Presidente nota che, così come le ispezioni, anche le richieste di informazioni determinano un'interferenza nel diritto al rispetto della vita privata e delle comunicazioni tutelato dagli artt. 7 e 52, par. 1, della Carta dei diritti fondamentali dell'UE, e dell'art. 8, par. 2, della Convenzione Europea per i Diritti dell'Uomo⁷. L'interferenza non si traduce in una violazione dei suddetti diritti fondamentali qualora essa possa essere considerata effettivamente giustificata ai sensi di un'adeguata base giuridica. Nella fattispecie, tale base giuridica non può essere individuata nell'art. 18, par. 3, del regolamento CE 1/2003 in quanto, come osservato in precedenza, la richiesta di informazioni può comportare l'acquisizione di informazioni non necessarie ai fini del procedimento⁸.

La possibilità che la richiesta di informazioni comporti l'acquisizione di informazioni irrilevanti deve essere valutata con particolare attenzione con riferimento ai dati personali di natura medica, configuranti informazioni di carattere "sensibile" secondo quanto previsto dal regolamento (UE) 2019/679 (GDPR)¹⁰ e dal regolamento (UE) 2018/1725 relativo al trattamento dei dati personali da parte delle istituzioni UE¹¹. Se

⁶ Ivi, §§ 47-53. Nel caso *Nexans France SAS*, la Corte di giustizia dell'Unione europea ha ritenuto che, nell'ambito di indagini antitrust, la Commissione possa copiare dati senza un loro esame preliminare a condizione che successivamente sia verificato, «nel rigoroso rispetto dei diritti della difesa dell'impresa interessata, che tali dati siano pertinenti per l'oggetto degli accertamenti, prima di inserire nel fascicolo i documenti ritenuti pertinenti al riguardo e di cancellare gli altri dati copiati». (vd. CGUE, C-606/18 P, *Nexans France SAS* (2020), § 64. Nel caso *Prysmian SpA*, in relazione ai poteri di ispezione ai sensi dell'art. 20 del Regolamento (CE) n. 1/2003, la Corte di giustizia ha confermato che il diritto della Commissione di procedere alla realizzazione di copie dei documenti acquisiti in sede di ispezione «non pregiudica né le garanzie procedurali previste dal regolamento n. 1/2003 né gli altri diritti dell'impresa che è oggetto di ispezione, a condizione che la Commissione, dopo aver completato il suo esame, inserisca nel fascicolo soltanto documenti che sono pertinenti tenuto conto dell'oggetto dell'ispezione» (vd. CGUE, C-601/18 P, *Prysmian SpA* (2020), § 59).

⁷ Vd. ordinanza, § 57. Nel caso *Xi c. Commissione*, il Tribunale ha riconosciuto che il trattamento dei dati medici richiede un esame particolarmente rigoroso in considerazione della loro natura estremamente intima e sensibile. La riservatezza delle informazioni sanitarie configura l'oggetto di uno dei diritti fondamentali protetti dall'ordinamento giuridico dell'Unione (vd. CGUE, T-528/18, *Xi c. Commissione* (2019), § 67).

⁸ Vd. ordinanza, § 61.

⁹ Ivi, §§ 63-65.

¹⁰ Vd. in particolare l'art. 9 del [regolamento \(UE\) 2016/679](#) del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati) in *GUUE* L 119 del 4.5.2016, 1-88.

¹¹ Vd. in particolare l'art. 10 del [regolamento \(UE\) 2018/1725](#) del Parlamento europeo e del Consiglio, del 23 ottobre 2018, sulla tutela delle persone fisiche in relazione al trattamento dei dati personali da parte delle istituzioni, degli organi e degli organismi dell'Unione e sulla libera circolazione di tali dati, e che abroga il regolamento (CE) n. 45/2001 e la decisione n. 1247/2002/CE in *GUUE* L 295 del 21.11.2018, 39-98. In attuazione del regolamento (UE) 2018/1725, con la [decisione \(EU\) 2018/1927](#) (in *GUUE* L 313 del 10.12.2018, 39-44) la Commissione europea ha stabilito norme interne relative al trattamento dei dati personali da parte della Commissione europea nel settore della concorrenza in relazione alla comunicazione di informazioni agli interessati e alla limitazione di determinati diritti.

risulta inevitabile che le indagini antitrust possano includere dati personali¹², il trattamento di questi ultimi deve sempre avvenire nel rispetto dei principi di necessità e proporzionalità in rapporto ai poteri conferiti alla Commissione europea. Tali principi non risultano *prima facie* rispettati nel caso di specie, nel quale la richiesta di informazioni ai sensi dell'art. 18, par. 3, del regolamento (CE) n. 1/2003 è stata formulata senza prevedere (i) né un metodo di verifica della rilevanza dei documenti selezionati mediante parole chiave, (ii) né misure atte a tutelare i dati personali dei soggetti interessati. Sul punto, conclude il Presidente che «*the fact that the Commission uses, for the purposes of its investigations, search methods which inevitably require the processing of personal data does not mean that it is not required to take account of the sensitivity of some of that data*»¹³.

Sotto il profilo del *periculum in mora*, il Presidente osserva che la circostanza che la Commissione sia tenuta - nell'ambito del procedimento antitrust - a stabilire la rilevanza dei documenti contenenti dati personali non comporta di per sé l'insorgere di un danno grave ed irreparabile per le persone interessate¹⁴, considerato che le informazioni in questione non verranno rese pubbliche ed i funzionari sono vincolati da stringenti obblighi di riservatezza¹⁵.

Tale assunto generale circa l'inconfigurabilità di un danno rilevante ai fini della sospensione della decisione non vale, tuttavia, per la particolare categoria di dati personali "sensibili" di natura medica: per tale categoria di dati il solo fatto di allargare la cerchia dei soggetti che ne sono a conoscenza è idoneo a causare un danno grave ed irreparabile per le persone interessate, che non potrebbero al riguardo trarre beneficio dall'eventuale successivo annullamento della decisione finale adottata dalla Commissione ai sensi del regolamento (CE) n. 1/2003¹⁶.

Per evitare dunque che tale danno si concretizzi, il Presidente dispone in conclusione la sospensione della decisione controversa fintanto che venga attuata una procedura *ad hoc* così articolata¹⁷:

l'impresa deve selezionare e trasmettere su supporto elettronico separato i documenti contenenti i dati personali sensibili;

tali documenti devono essere quindi inclusi in una *virtual data room* accessibile ad un numero ristretto di funzionari della Commissione e ad un uguale numero di avvocati della stessa impresa;

¹² European Data Protection Supervisor (EDPS), lettera del 22 ottobre 2018 (WW/OL/sn/D (2018) 2422 C 2018-0632) ove si rileva che, sebbene «*Commission investigations and enforcement activities in the competition field target undertakings or Member States which are subject to the competition rules of the Treaty, and not natural persons as such, [...] during competition investigations inevitably also personal data are being processed*».

¹³ Vd. ordinanza, § 67.

¹⁴ Ivi, § 84.

¹⁵ Ivi, §§ 77-81.

¹⁶ Ivi, §§ 86-87. Nel caso *Commissione c. Akzo*, la Corte ha stabilito che qualora una decisione recante una richiesta di informazioni sia dichiarata nulla nel merito, la Commissione europea è tenuta a rimuovere dal fascicolo del procedimento i documenti acquisiti illegittimamente e a non utilizzarli quali prove nell'ambito del procedimento (vd. CGUE, C-7/04 P(R), *Commissione c. Akzo* (2004), §§ 37-39). L'estromissione dei documenti dal fascicolo non è in grado di sanare l'eventuale diffusione dei dati personali sensibili per i quali, secondo il Presidente del Tribunale, il pregiudizio si concretizza con l'ampliamento dei soggetti a conoscenza di tali informazioni.

¹⁷ Vd. ordinanza, §§ 119-121.

i funzionari devono esaminare e selezionare i documenti in questione ritenuti rilevanti, dando la possibilità agli avvocati dell'impresa di esprimere osservazioni prima dell'inserimento di ciascun documento nel fascicolo del procedimento; in caso di disaccordo, il documento non potrà essere inserito nel fascicolo e gli avvocati dell'impresa dovranno esprimere le loro osservazioni al riguardo. In caso di persistente disaccordo, l'impresa potrà chiedere l'intervento del Direttore dell'Informazione, Comunicazione e Media presso la Direzione Generale Concorrenza, il quale dovrà dirimere la controversia.

2. Il quadro normativo UE

Il rapporto tra i poteri di indagini delle istituzioni UE e la tutela dei dati personali è stato nel 2018 oggetto di una lettera di chiarimento da parte dell'EDPS¹⁸, il quale è stato chiamato a precisare i limiti ai quali gli operatori economici devono attenersi, a seguito dell'avvento del GDPR, nel rispondere alle richieste di informazioni provenienti dalle pubbliche autorità europee. La lettera di chiarimento si concentra su alcuni quesiti specifici relativi in particolare:

alla possibilità per gli operatori economici di rifiutare di fornire informazioni in quanto costituenti dati personali ai sensi del GDPR;

alla sussistenza di un obbligo legale ai sensi dell'art. 14 GDPR, a carico degli stessi operatori economici, di informare le persone fisiche interessate qualora dati personali ad esse relativi siano comunicati ad un'autorità UE.

In risposta a tali quesiti, per quanto concerne specificamente la Direzione Generale Concorrenza della Commissione europea (DG Comp), l'EDPS rileva che nell'ambito dei poteri di indagine sull'attività delle imprese conferiti ai sensi dei Trattati e dei Regolamenti UE in materia di controllo delle concentrazioni, antitrust ed aiuti di Stato, la Commissione si trova necessariamente a trattare dati personali.

In tale contesto generale, l'EDPS rileva innanzitutto che il GDPR si pone in continuità con il quadro normativo di cui alla previgente direttiva 95/47/CE: da un lato, le istituzioni UE sono obbligate a rispettare elevati standard di tutela dei dati personali ai sensi sia del GDPR sia del regolamento (UE) 2018/1725, che ha abrogato il previgente regolamento (CE) 45/2001 sulla stessa materia; d'altro lato, ai sensi dell'art. 6, par. 1, lett. e), GDPR e dell'art. 5, par. 1, lett. a), del regolamento (UE) 2018/1725, le istituzioni UE provvedono legittimamente a raccogliere e trattare dati personali qualora agiscano sulla base ed entro i limiti dei poteri ad esse conferiti per legge¹⁹.

In risposta al primo quesito, dunque, l'EDPS riconosce che il GDPR non impedisce agli operatori economici di fornire informazioni contenenti dati personali (sia in adempimento ad un obbligo legale, sia su base volontaria), purché l'istituzione che riceve i

¹⁸ EDPS, lettera del 22 ottobre 2018, cit.

¹⁹ L'art. 5, par. 1, lett. a), del regolamento (UE) 2018/1725 stabilisce in particolare che «[i]l trattamento è lecito solo se e nella misura in cui ricorre almeno una delle seguenti condizioni: a) il trattamento è necessario per l'esecuzione di un compito svolto nell'interesse pubblico o nell'esercizio di pubblici poteri di cui sono investiti l'istituzione o l'organo dell'Unione».

dati agisca nell'ambito dei propri poteri e della propria sfera di competenze²⁰.

In risposta al secondo quesito, attinente all'obbligo degli operatori economici in qualità di titolari del trattamento di informare le persone interessate qualora i loro dati siano trasmessi a soggetti terzi c.d. "destinatari" ai sensi dell'art. 14, par. 1, lett. e), GDPR, l'EDPS evidenzia che «le autorità pubbliche che possono ricevere comunicazione di dati personali nell'ambito di una specifica indagine conformemente al diritto dell'Unione o degli Stati membri non sono considerate destinatari» per espresso disposto dell'art. 4, n. 9, GDPR²¹. Per l'effetto, gli operatori economici non hanno l'obbligo legale di informare le persone interessate circa la trasmissione dei loro dati personali alle istituzioni dell'UE nel caso in cui tali dati siano comunicati nell'ambito di uno specifico procedimento di indagine, svolto dalle istituzioni stesse ai sensi dei poteri conferiti dal diritto dell'UE²².

Stante l'oggetto specifico delle domande che ne costituiscono il fondamento, i chiarimenti dell'EDPS non entrano tuttavia nel merito delle particolari misure che devono essere adottate dalle istituzioni UE nell'ambito del procedimento, a tutela della riservatezza dei dati personali acquisiti. E proprio a questo riguardo interviene l'ordinanza del Presidente del Tribunale in commento, nella parte in cui – per colmare tale "vuoto" – va a definire un'essenziale procedura *ad hoc* per il trattamento dei dati personali sanitari in conformità sia all'art. 18, par. 3, del regolamento 1/2003 CE, sia ai diritti fondamentali previsti dagli artt. 7 e 52, par. 1, della Carta dei diritti fondamentali dell'UE e dell'art. 8, par. 2, della Convenzione europea per i diritti dell'uomo.

3. Metodologie di e-discovery e *predictive coding*: uno sguardo all'esperienza USA e UK

Volgendo lo sguardo al di là dell'Atlantico e della Manica, possiamo individuare paradigmi più strutturati per l'impiego generalizzato delle tecniche di e-discovery, nell'ambito dei procedimenti amministrativi e giudiziali, fondate sull'utilizzo del *predictive coding* (o *technology-assisted review*, TAR).

In ambito giuridico, con tale nozione si intendono i sistemi software che, grazie all'impiego dell'intelligenza artificiale in combinazione con quella umana, sono in grado di automatizzare parzialmente le procedure di e-discovery, riducendo il numero di docu-

²⁰ In risposta al primo quesito, letteralmente l'EDPS precisa in conclusione che: «[t]he GDPR does not prevent the submission of information containing personal data to EU institutions, either in response to a legal obligation to do so or on a voluntary basis, as long as you act within your powers and sphere of competences».

²¹ Al riguardo si veda il considerando 31 del GDPR: «Le autorità pubbliche a cui i dati personali sono comunicati conformemente a un obbligo legale ai fini dell'esercizio della loro missione istituzionale, quali autorità fiscali e doganali, unità di indagine finanziaria, autorità amministrative indipendenti o autorità dei mercati finanziari, responsabili della regolamentazione e della vigilanza dei mercati dei valori mobiliari, non dovrebbero essere considerate destinatari qualora ricevano dati personali che sono necessari per svolgere una specifica indagine nell'interesse generale, conformemente al diritto dell'Unione o degli Stati membri. [...]».

²² In risposta al secondo quesito, l'EDPS precisa infine che: «[e]conomic operators do not have the legal obligation to inform people about the disclosure of their personal data to EU institutions in case such data is submitted to your services with a view to carrying out a particular inquiry within their powers under Union law».

menti irrilevanti oggetto di “revisione lineare”, documento per documento, da parte dei team di legali. In tali sistemi, si chiede al software dapprima di esaminare un campione di documenti classificati manualmente dal team di legali esperti. Avendo così appreso i criteri di classificazione, il software passa quindi ad analizzare nuovi set di documenti per determinarne la rilevanza in relazione alla richiesta di informazioni originaria. Campioni dei documenti classificati dal software vengono a tal punto riesaminati dal team di legali, al fine di stabilire il tasso di accuratezza dei risultati. Laddove questi ultimi siano insoddisfacenti, si richiede al software di ripetere l’analisi apprendendo dagli errori precedentemente commessi, sino a restituire risultati attendibili.

Negli USA il *predictive coding* ha fatto il proprio formale ingresso nelle aule giudiziarie a partire dall’*opinion* del giudice A. J. Peck nel caso *Da Silva Moore v. Publicis Groupe et al.*²³. Citando un proprio precedente scritto al riguardo, il giudice Peck ha definito in quell’occasione le tecniche di e-discovery assistite mediante l’uso dell’intelligenza artificiale come segue:

*«By computer-assisted coding, I mean tools (different vendors use different names) that use sophisticated algorithms to enable the computer to determine relevance, based on interaction with (i.e., training by) a human reviewer. Unlike manual review, where the review is done by the most junior staff, computer-assisted coding involves a senior partner (or [small] team) who review and code a “seed set” of documents. The computer identifies properties of those documents that it uses to code other documents. As the senior reviewer continues to code more sample documents, the computer predicts the reviewer’s coding (Or, the computer codes some documents and asks the senior reviewer for feedback.). When the system’s predictions and the reviewer’s coding sufficiently coincide, the system has learned enough to make confident predictions for the remaining documents.»*²⁴.

Quanto alle metodologie di *predictive coding*, il giudice Peck raccomanda l’utilizzo del modello contenuto nella Sedona Cooperation Proclamation²⁵, incentrato non tanto sulla soluzione del dilemma della “Black Box” (relativo alla (in)conoscibilità delle regole di giudizio effettivamente applicate dal software), quanto piuttosto sulla verificabilità e attendibilità dei risultati in relazione ai documenti forniti al sistema di intelligenza artificiale ai fini dell’apprendimento delle regole di giudizio ed alla successiva revisione dei risultati così ottenuti²⁶. L’applicazione del *predictive coding* richiede, dunque, l’elaborazione e soprattutto la condivisione tra le parti in causa della metodologia impiegata per la selezione dei documenti rilevanti, da definirsi con l’intervento del giudice ove necessario per la composizione di eventuali punti di disaccordo.

Negli USA il *predictive coding* è applicato anche dalla Direzione Antitrust del DoJ

²³ *Da Silva Moore v. Publicis Groupe et al.*, No. 1:2011cv01279 - Document 96 (S.D.N.Y. 2012). Si veda anche A.J. Peck, *Search, forward. Will manual document review and keyword searches be replaced by computer-assisted coding?*, in Law Technology News, Oct. 2011; W.W. Belt - D.R. Kiker - D.E. Shetterly, *Technology-Assisted Document Review: Is it Defensible?*, in *The Richmond Journal of Law & Technology*, XVIII, 3, 2012, 10 ss.

²⁴ *Da Silva Moore v. Publicis Groupe*, cit., 3-4.

²⁵ *The Sedona Conference Cooperation Proclamation*, in *thesedonaconference.org*, 2008.

²⁶ *Da Silva Moore v. Publicis Groupe*, cit., 4-5.

nell'ambito dei procedimenti di e-discovery²⁷, secondo linee guida metodologiche che possono essere utilizzate dalle imprese che ne facciano richiesta nell'ambito dei procedimenti antitrust in risposta alla c.d. Second Request. Ove l'impresa interessata dal procedimento intenda avvalersi del *predictive coding*, dovrà sottoscrivere con il DoJ un apposito accordo nel quale vengono indicati, in particolare, il software utilizzato ed i relativi "recall rate" e margini di errore, nonché i criteri da applicarsi nelle fasi di training del software, controllo qualitativo e validazione dei risultati²⁸.

Su indicazione del DoJ, le imprese interessate devono attenersi alle seguenti istruzioni in merito al trattamento dei dati personali ed in particolare di tipo sanitario contenuti nella documentazione interessata dal procedimento:

«Personally identifiable information (PII), medical information, or other protected information. If the production will include PII, information protected by the Health Insurance Portability and Accountability Act (HIPAA), or other information that is protected by statute or regulation, notify staff as soon as possible, but in any event, before producing it. The Division is open to discussing modifications that will permit the Company to withhold this kind of information entirely. If such a modification is not practicable, notice will allow the Division to institute appropriate protections and procedures to protect this sensitive information»²⁹.

In applicazione di tali linee guida, dunque, è onere dell'impresa notificare al DoJ la presenza di dati personali nei documenti prima di produrli nel procedimento. Ove possibile, d'accordo con il DoJ, si valuterà in prima battuta l'integrale omissione dei dati personali; in subordine, ove ciò non risulti possibile, dovranno essere applicate apposite misure di tutela dei dati.

Nel Regno Unito, nel 2016 la High Court of England and Wales ha ammesso l'utilizzo dei sistemi di *predictive coding* nel caso *Pyrrho Investments*³⁰. Nella fattispecie, il ricorso al *predictive coding* è stato ritenuto appropriato in quanto necessario all'esame di circa 3 milioni di documenti, ossia di una mole tale di documenti per la quale il software è in grado di restituire risultati maggiormente corretti ed omogenei, in relazione ai criteri di classificazione documentale forniti nella fase di training da parte di un avvocato esperto, rispetto ad un'analisi esclusivamente manuale. L'impiego del *predictive coding* è stato ritenuto inoltre in grado di determinare una riduzione dei costi del procedimento, rispetto alle spese derivanti dallo svolgimento della stessa attività da parte di un team di legali.

Più recentemente, ossia dal 2020, la Competition and Markets Authority (CMA) ha iniziato a testare sistemi di *predictive coding* con l'obiettivo di ridurre i tempi del proce-

²⁷ T. Greer, *Electronic Discovery at the Antitrust Division: An Update*, in *justice.gov*, 25 giugno 2015; Id., *Avoiding E-Discovery Accidents & Responding to Inevitable Emergencies: A Perspective from the Antitrust Division*, in *justice.gov*, marzo 2017.

²⁸ U.S. DoJ, *Predictive Coding Model Agreement*, in *justice.gov*, ultima versione disponibile aggiornata al 25 settembre 2018.

²⁹ U.S. DoJ, *Form of Production of ESI Documents in Response to the Second Request*, in *justice.gov*.

³⁰ *Pyrrho Investments Limited v MWB Property Limited* [2016] EWHC 256 (Ch). In proposito, A. Cunningham - A.D. James - P. Taylor - B. Tether, *Disruptive Technologies & Legal Service Provision in the UK: A Preliminary Study*, 7 dicembre 2018.

dimento e di impiegare in modo più efficiente le proprie risorse interne³¹. Da quanto reso noto dalla CMA, il *predictive coding* sembrerebbe essere utilizzato nelle procedure interne, a vantaggio degli uffici dell’Autorità, allo scopo di esaminare in via prioritaria quei documenti - tra quelli acquisiti dalla CMA - individuati dal software come «relevant for early review». I contenuti e gli esiti di tale sperimentazione dovrebbero auspicabilmente essere condivisi in futuro, allo scopo di costruire metodologie appropriate e trasparenti di *predictive coding* nell’interesse non solo dell’efficienza del procedimento ma, soprattutto, dei diritti fondamentali di difesa delle imprese e delle persone fisiche coinvolte.

³¹ S. Nicols, *Predictive coding: how technology could help to streamline cases*, CMA Blog, 24 luglio 2020, in competitionandmarkets.blog.gov.uk. Il *predictive coding* risulta essere in fase di sperimentazione a partire da alcuni casi del 2020, relativi ad accordi anticompetitivi di rivendita online degli strumenti musicali, nei quali la CMA aveva acquisito oltre 10 milioni di documenti.