Provvedimento del 2 novembre 2024 [10085455]

VEDI ANCHE

- Comunicato del 20 dicembre 2024

- Comunicato del 13 aprile 2023

- Comunicato del 12 aprile 2023

- Provvedimento dell'11 aprile 2023

- Comunicato dell'8 aprile 2023

- Comunicato del 6 aprile 2023

- Comunicato del 4 aprile 2023

Comunicato del 31 marzo 2023

- Provvedimento del 30 marzo 2023

[doc. web n. 10085455]

Provvedimento del 2 novembre 2024

Registro dei provvedimenti
n. 755 del 2 novembre 2024

IL GARANTE PER LA PROTEZIONE DEI DATI PERSONALI

NELLA riunione odierna, alla quale hanno preso parte il prof. Pasquale Stanzione, presidente, la prof.ssa Ginevra Cerrina Feroni, vicepresidente, il dott. Agostino Ghiglia e l'avv. Guido Scorza, componenti, e il cons. Fabio Mattei, segretario generale;

VISTO il 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 personali, di seguito "Regolamento");

VISTO il Codice in materia di protezione dei dati personali (d.lgs. 30 giugno 2003, n. 196), come modificato dal d.lgs. 10 agosto 2018, n. 101, recante disposizioni per l'adeguamento dell'ordinamento nazionale al citato Regolamento (di seguito "Codice");

VISTO il Regolamento n. 1/2019 concernente le procedure interne aventi rilevanza esterna, finalizzate allo svolgimento dei compiti e all'esercizio dei poteri demandati al Garante per la protezione dei dati personali, approvato con deliberazione n. 98 del 4 aprile 2019, pubblicato in G.U. n. 106 dell'8 maggio 2019 e in www.gpdp.it, doc. web n. 9107633 (di seguito "Regolamento del Garante n. 1/2019");

VISTA la documentazione in atti;

VISTE le osservazioni formulate dal segretario generale ai sensi dell'art. 15 del Regolamento del Garante n. 1/2000;

RELATORE la prof.ssa Ginevra Cerrina Feroni;

1. INTRODUZIONE

Il procedimento trae origine da una attività istruttoria avviata d'ufficio a seguito della pubblicazione di notizie stampa relative ad alcune problematiche tecniche (bug) occorse il giorno 20 marzo 2023 al servizio ChatGPT offerto e gestito dalla società statunitense OpenAI OpCo, LLC (di seguito anche la "Società" o "OpenAI").

ChatGPT, acronimo di Chat Generative Pre-trained Transformer, è un chatbot basato su di un sistema di intelligenza artificiale generativa, che genera output testuali sulla base di input (cd. prompt) testuali.

Ai fini di questo provvedimento, per "intelligenza artificiale generativa" si intende il campo dell'intelligenza artificiale che si concentra sulla creazione di contenuti nuovi e originali utilizzando algoritmi prevalentemente di tipo neurale e per "rete neurale" si intende un modello computazionale standard applicabile nei contesti più diversificati che permette il riconoscimento di oggetti, forme o pattern all'interno di un dato o un insieme di dati (ad esempio, un volto umano in una fotografia). Gli algoritmi di intelligenza artificiale generativa sono impiegati in una vasta gamma di applicazioni, tra cui il riconoscimento e la generazione di immagini, di tracce vocali o musicali, di testi e di video.

La versione del modello di linguaggio di grandi dimensioni (Large Language Model – LLM), all'epoca dei fatti sottesa al servizio gratuito di ChatGPT è GPT 3.5, quella alla versione a pagamento del servizio è GPT 4.

Ai fini di questo provvedimento per "Large Language Model" si intende un modello probabilistico di un linguaggio naturale, come la lingua inglese o italiana, che si basa sulla constatazione che tutti i linguaggi naturali sono fortemente ridondanti e correlati, e per "Generative Pre-trained Transformer (GPT)" si intende un algoritmo di intelligenza artificiale generativa che si fonda sull'uso di un particolare modello computazionale detto Transformer, che appare più efficiente e più versatile del classico modello neurale nei casi in cui la struttura dei dati in ingresso sia di tipo sequenziale (come nei Large Language Models) e l'obiettivo della generazione sia la previsione del successivo elemento della sequenza sulla base dell'osservazione di tutti gli elementi precedenti. L'impiego di particolari modelli computazionali (Transformer) e di particolari rappresentazioni numeriche delle unità linguistiche (Embedding) consente di costruire una rete molto fitta e molto estesa di correlazioni semantiche tra unità linguistiche in un testo, rendendo la generazione automatizzata di nuovi testi praticamente indistinguibile dal testo prodotto in forma creativa da un essere umano che abbia letto lo stesso corpus di testi comprendendone il senso.

Le notizie stampa hanno reso noto che, a causa di un bug, sulla pagina principale del servizio ChatGPT, l'utente visualizzava la cronologia dei titoli delle chat di altri utilizzatori del servizio anziché le proprie.

In seguito la Società ha pubblicamente confermato l'accaduto e ha precisato che i dati coinvolti nella problematica tecnica che avrebbero potuto essere visualizzati da utenti diversi dagli interessati erano il nome, il cognome, l'indirizzo e-mail, nonché le ultime quattro cifre e la scadenza della carta di credito utilizzata per il pagamento del servizio ChatGPT Plus (la versione a pagamento del servizio).

Alla luce della notizia di tale data breach il Garante ha avviato un'istruttoria ex officio rilevando che il trattamento dei dati personali da parte di OpenAI nell'ambito del servizio ChatGPT potesse dare luogo ad una violazione della normativa in materia di dati personali con particolare riferimento all'assenza di una idonea informativa per gli utenti e per tutti gli interessati i cui dati fossero stati raccolti da OpenAI e trattati nell'ambito del servizio ChatGPT; all'assenza di una base giuridica per il trattamento dei dati a fini di addestramento degli algoritmi sottesi al funzionamento della piattaforma; alla non corrispondenza di alcune delle informazioni fornite da ChatGPT al dato reale e la conseguente inesattezza dei dati personali oggetto delle attività di trattamento del titolare; all'assenza di un qualsivoglia filtro per la verifica dell'età degli utenti, sebbene il servizio fosse rivolto ad utenti maggiori di 13 anni, con il conseguente rischio di esposizione dei minori a risposte inidonee rispetto al loro grado di sviluppo ed autoconsapevolezza.

In tale quadro, il 30 marzo 2023, il Presidente dell'Autorità ha adottato nei confronti di OpenAI, ex art. 5, comma 8, del regolamento del Garante n. 1/2000, un provvedimento d'urgenza (n. 112/2023, prot. n. 54718/23) di limitazione provvisoria del trattamento dei dati personali degli interessati stabiliti nel territorio italiano, ai sensi dell'art. 58, par. 2, lett. f), del Regolamento.

Successivamente, con il provvedimento n. 114 dell'11 aprile 2023, l'Autorità ha deliberato la sospensione del provvedimento n. 112/2023 di limitazione provvisoria a condizione che il titolare, ex art. 58, par. 2, lett. d) del Regolamento, adottasse misure idonee a garantire che le attività di trattamento dei dati personali nell'ambito del servizio ChatGPT avvenissero in modo conforme alla normativa in materia di protezione dei dati personali. Segnatamente l'Autorità ha ingiunto al titolare di:

1. predisporre e pubblicare nel proprio sito web un'informativa che, nei termini e con le modalità di cui all'art. 12 del Regolamento, rendesse noti agli interessati, anche diversi dagli utenti del servizio ChatGPT, le attività di raccolta e di trattamento dei loro dati ai fini dell'addestramento degli algoritmi, le modalità del trattamento, la logica alla base del trattamento necessario al funzionamento del servizio, i diritti loro spettanti in qualità di interessati e ogni altra informazione prevista dal Regolamento;

2. mettere a disposizione, nel proprio sito web, agli interessati, anche diversi dagli utenti del servizio, che si fossero collegati dall'Italia, uno strumento attraverso il quale poter esercitare il diritto di opposizione rispetto ai trattamenti dei propri dati personali, ottenuti da terzi, svolti dalla società ai fini dell'addestramento degli algoritmi e dell'erogazione del servizio;

3. mettere a disposizione, nel proprio sito web, agli interessati, anche diversi dagli utenti del servizio, che si fossero collegati dall'Italia, uno strumento attraverso il quale poter chiedere e ottenere la correzione dei loro dati personali se trattati in maniera inesatta nella generazione dei contenuti oppure, laddove tale misura fosse risultata impossibile allo stato della tecnica, di procedere alla cancellazione dei propri dati personali;

4. inserire un link all'informativa rivolta agli utenti del servizio nel flusso di registrazione in una posizione che ne consentisse la lettura prima di procedere alla registrazione, attraverso modalità tali da consentire a tutti gli utenti che si fossero collegati dall'Italia, ivi inclusi quelli già registrati, al primo accesso successivo all'eventuale riattivazione del servizio, di prendere visione di tale informativa;

5. modificare la base giuridica del trattamento dei dati personali degli utenti ai fini dell'addestramento degli algoritmi, eliminando ogni riferimento al contratto e assumendo come base giuridica del trattamento il consenso o il legittimo interesse in relazione alle valutazioni di competenza della società in una logica di accountability;

6. mettere a disposizione, nel proprio sito Internet, agli utenti del servizio che si fossero

collegati dall'Italia, uno strumento facilmente accessibile attraverso il quale esercitare il diritto di opposizione al trattamento dei propri dati acquisiti in sede di utilizzo del servizio per l'addestramento degli algoritmi qualora la base giuridica prescelta ai sensi del punto 5 che precede fosse stata individuata nel legittimo interesse;

7. in sede di eventuale riattivazione del servizio dall'Italia, inserire la richiesta, a tutti gli utenti collegati dall'Italia, ivi inclusi quelli già registrati, di superare, in sede di primo accesso, un age gate tale da escludere, sulla base dell'età dichiarata, gli utenti minorenni;

8. sottoporre al Garante, entro il 31 maggio 2023, un piano per l'adozione di strumenti di age verification idoneo ad escludere l'accesso al servizio agli utenti infra tredicenni e a quelli minorenni in assenza di un'espressa manifestazione di volontà da parte dei titolari della responsabilità genitoriale, da implementare, al più tardi, dal 30 settembre 2023;

9. promuovere, entro il 15 maggio 2023, una campagna di informazione, di natura non promozionale, su tutti i principali mezzi di comunicazione di massa italiani (radio, televisione, giornali e Internet), previo accordo con il Garante circa i contenuti, allo scopo di informare le persone dell'avvenuta probabile raccolta dei loro dati personali ai fini dell'addestramento degli algoritmi, dell'avvenuta pubblicazione sul sito internet della Società di un'apposita informativa di dettaglio e della messa a disposizione, sempre sul sito internet della Società, di uno strumento che consentisse agli interessati di chiedere ed ottenere la cancellazione dei propri dati personali.

Il Garante ha indicato termini differenziati per l'attuazione di ciascuna delle sopra indicate prescrizioni, stabilendo il termine ultimo del 30 aprile 2023 per l'adempimento delle prescrizioni di cui ai punti da 1 a 7; il termine del 31 maggio 2023 per la prescrizione di cui al punto 8 ed il termine del 15 maggio 2023 per l'adempimento della prescrizione di cui al punto 9, riservandosi ogni ulteriore intervento, anche di carattere urgente e temporaneo, nel caso di inidonea o insufficiente attuazione delle misure prescritte.

2. LE RISPOSTE DI OPENAI AI PROVVEDIMENTI NN. 112/2023 E 114/2023 E LE OSSERVAZIONI DEL GARANTE

Con nota del 28 aprile 2023 (prot. n. 69713/23) OpenAI ha comunicato di aver adempiuto alle richieste di cui ai punti da 1 a 7 e, in particolare:

1. di aver adottato e pubblicato nelle sezioni privacy policy ed help center del sito una privacy policy relativa al trattamento per finalità di training dei modelli;

2. di aver implementato, in aggiunta ad un indirizzo e-mail dedicato per la presentazione delle istanze degli interessati, uno strumento per consentire agli utenti europei di esercitare i diritti di opposizione e di cancellazione compilando un apposito form reso disponibile nella privacy policy e nella notice del sito;

3. in relazione al diritto di rettifica di continuare gli approfondimenti e le ricerche in merito, consentendo, ad ogni modo, agli interessati di rivolgere le loro istanze al titolare inviando un'e-mail dedicata oppure utilizzando il form di cui al punto precedente;

4. di aver inserito il link alla privacy policy nella home page del sito, nella pagina di log-in e nella pagina di registrazione in una posizione tale da consentirne la lettura prima della conclusione della procedura di registrazione e che agli utenti già registrati sarebbe stata mostrata una finestra contenente i link alla privacy policy ed all'help center, unitamente alla richiesta di conferma della loro età;

5. di aver individuato il legittimo interesse di cui all'art. 6, par.1, lett. f), del Regolamento

come base giuridica per il trattamento dei dati personali ai fini di addestramento degli algoritmi;

6. di aver reso disponibile - attraverso i ToS (Terms of Service), la privacy policy e la sezione Data usage for consumer services FAQ - un form per consentire agli utenti di esercitare un opt-out rispetto all'utilizzo dei propri dati personali per fini di addestramento degli algoritmi;

7. di aver implementato un age gate che prevede: a) per i nuovi utenti, di fornire la data di nascita, in modo tale che, da un lato, sia impedita la creazione dell'account qualora venga indicata un'età inferiore ai 13 anni e, dall'altro, si ottenga la conferma da parte dell'interessato di età compresa tra i 13 ed i 17 anni dell'avvenuta raccolta del consenso del titolare della responsabilità genitoriale alla creazione dell'account; b) per gli utenti già registrati, di dichiarare una età superiore ai 13 anni e, se minori di anni 18, di aver raccolto il consenso del titolare della responsabilità genitoriale per accedere ai servizi ChatGPT.

Il titolare ha chiarito altresì di aver integrato la privacy policy indicando espressamente, per i minori di età compresa tra i 13 ed i 18 anni, la necessità del consenso di un genitore ovvero del titolare della responsabilità genitoriale per accedere ai servizi ChatGPT, come già indicato nei ToS.

Con successiva nota del 15 maggio 2023 (prot. n. 78218/23) OpenAI ha comunicato di aver adempiuto alla richiesta di cui al punto 9 del provvedimento n.114/23 attraverso:

a) interviste sui quotidiani: in particolare, "La Repubblica", nell'inserto Affari&Finanza del 15 maggio, ha pubblicato un'intervista a Mira Murati, CTO della Società, volta anche ad evidenziare le misure che gli interessati possono adottare laddove non vogliano che i loro dati vengano utilizzati per addestrare i modelli sottesi ai servizi ChatGPT;

b) avvisi sui quotidiani: in particolare, è stato realizzato uno spazio pubblicitario dedicato, sulle edizioni cartacee di lunedì 15 maggio de "La Repubblica" e de "La Stampa", per pubblicare materiale educativo relativo al servizio ChatGPT;

c) una pagina diretta agli utenti nel sito web di OpenAI;

d) un video didattico da realizzare in collaborazione con l'Autorità: in particolare, la Società ha rappresentato al Garante di aver avviato la realizzazione di un video relativo ai rilievi indicati nel punto 9 di cui al provvedimento n. 114/23 e ha manifestato la volontà di adottare lo stesso in collaborazione con il Garante rendendolo pubblico tramite i canali istituzionali dell'Autorità.

L'Autorità con nota del 18 maggio 2023 (prot. n. 79806/23) ha espresso forte contrarietà rispetto alla campagna mediatica, come sopra prospettata, in quanto realizzata senza alcuna preventiva informazione ed accordo con il Garante in merito ai termini ed alle condizioni della stessa, contrariamente a quanto prescritto nel provvedimento n. 114/2023. Con la stessa nota l'Autorità si è riservata ogni valutazione rispetto al puntuale adempimento delle prescrizioni impartite con il richiamato provvedimento ed ha invitato la Società a presentare senza ritardo e, comunque, non oltre il 19 maggio, un progetto di campagna di comunicazione in linea con la citata prescrizione e con le osservazioni rappresentate, idonea a raggiungere il medesimo pubblico i cui dati personali siano stati, verosimilmente, trattati.

La Società, a seguito di una richiesta di proroga, ha risposto con nota del 23 maggio 2023 (prot. n. 82299/23) con cui, da una parte, ha richiamato le iniziative già illustrate con la nota precedente finalizzate ad adempiere alle sopracitate prescrizioni dell'Autorità e, dall'altra, sempre in un'ottica di adeguamento alle indicazioni ricevute, ha rappresentato la volontà di volersi impegnare al fine di organizzare con emittenti televisive e radiofoniche nazionali la diffusione di messaggi informativi

inerenti al servizio ChatGPT. In particolare, la Società ha rappresentato che tali messaggi avrebbero:

- illustrato le modalità di addestramento di ChatGPT;

- chiarito che il training comporta anche il trattamento di dati personali;

- fornito indicazioni agli interessati sulle modalità per reperire maggiori informazioni ed esercitare i propri diritti ai sensi del Regolamento.

In riscontro, l'Autorità, con nota del 22 giugno 2023 (prot. n. 97898/23), ha precisato di non essere in grado, alla luce delle informazioni fornite, di valutare l'impatto della campagna con particolare riferimento alle iniziative che la Società avrebbe avuto intenzione di lanciare via social, mancando qualsivoglia elemento a ciò utile, e ritenendo le prospettate attività informative da svolgere mediante giornali e televisioni al di sotto delle aspettative e difficilmente idonee a raggiungere il pubblico sperato. Il Garante ha altresì espresso parere contrario all'utilizzazione del proprio logo nella menzionata campagna comunicazionale.

Con nota del 31 maggio 2023 (prot. n. 86958/23) OpenAI ha trasmesso una comunicazione relativa al sistema di age verification di cui al punto 8 del provvedimento n. 114/2023, prospettando, sulla base dello stato dell'arte, alcune possibili soluzioni di verifica dell'età, ed evidenziando la mancanza, a livello europeo, di una posizione condivisa in materia. In particolare, la Società ha proposto le seguenti soluzioni: i) la verifica da parte di un soggetto terzo; ii) l'impiego del sistema pubblico italiano di identità digitale (SPID); iii) l'impiego di tecnologie di intelligenza artificiale; iv) la verifica della carta di credito (in aggiunta ad altro sistema); v) l'impiego di metodi di conferma (ad es. domande specifiche) del consenso parentale.

A tale riguardo, il Garante, con la già citata nota del 22 giugno 2023, ha precisato che, ferma restando la possibilità di un costruttivo confronto, anche futuro, sul tema, spettasse alla Società, entro il mese di settembre, in una logica di accountability, individuare una o più soluzioni ritenute utili allo scopo e all'adempimento della prescrizione impartita con il citato provvedimento n. 114/2023.

OpenAI, con nota del 29 settembre 2023 (prot. n. 8657/24), ha comunicato di aver deciso di affidare l'attività di age verification ad una terza parte, nella fattispecie la società certificata Yoti Ltd. (di seguito "Yoti"), la quale garantisce che nell'esecuzione della verifica dell'età nessun dato viene fornito ad OpenAI. In particolare, la Società ha rappresentato di voler implementare le seguenti soluzioni di verifica dell'età degli utenti:

1.Yoti App: gli utenti forniscono, una sola volta a Yoti un documento di identità rilasciato dal governo ed una fotografia con successiva verifica per l'accesso ad OpenAI mediante scansione di un codice QR dalla pagina web di OpenAI;

2. Age estimation: gli utenti effettuano un autoscatto (selfie) utilizzando l'app o il sito web di Yoti, che stima l'età dell'utente con un tasso di errore assoluto di 1,79 anni;

3. ID Scan: gli utenti scansionano un documento di identità ed effettuano un autoscatto; in questo caso Yoti controlla la corrispondenza tra i due elementi utilizzando un sistema di elaborazione automatico.

3. ATTIVITÀ ISTRUTTORIA

Parallelamente alla gestione del provvedimento cautelare, l'Autorità ha proceduto con l'acquisizione degli elementi ritenuti necessari per lo svolgimento dell'istruttoria tramite due richieste di informazione, ai sensi degli articoli 58, par. 1, lett. e), del Regolamento e 157 del

Codice.

Con nota del 4 aprile 2023 (prot. n. 57229/23), il Garante ha inviato una richiesta di informazioni ad OpenAI chiedendo chiarimenti in merito: a) al funzionamento di ChatGPT; b) alle informazioni sul trattamento dei dati forniti agli interessati; c) alle misure adottate per impedire l'accesso al servizio ai soggetti minori di 13 anni e d) all'evento di data breach del 20 marzo 2023.

In seguito, con nota del 6 ottobre 2023 (prot. n. 137422/23), il Garante ha inviato alla Società una richiesta di informazioni integrativa, sia con riferimento specifico al sistema di age verification, sia relativamente al riscontro fornito da OpenAI alla nota precedente.

In risposta alla prima richiesta d'informazioni dell'Autorità, la Società con nota del 19 maggio 2023 (prot. n.80945/23), ha rappresentato:

a) in merito al funzionamento di ChatGPT:

- di aver addestrato il proprio modello con i dati provenienti da tre fonti primarie di informazioni: i) quelle pubblicamente disponibili su Internet, ii) quelle concesse in licenza da terzi e, infine, iii) quelle fornite dagli utenti o dagli "addestratori" incaricati dalla Società a tal fine. A tal proposito la Società ha allegato la valutazione di impatto (DPIA) datata 19 maggio 2023 (prima bozza del 24 febbraio 2023);

- di aver adottato misure per limitare la quantità di dati personali presenti nelle informazioni di addestramento del modello, escludendo i siti web che contengono grandi volumi di dati personali, non raccogliendo dati dal c.d. dark web e utilizzando, nella fase di fine-tuning del modello, anche Azure Cognitive Services per rimuovere le informazioni di carattere personale;

- di utilizzare diverse metodologie per la raccolta delle informazioni di addestramento: i) per quelle presenti in Internet, il c.d. crawling da parte direttamente di OpenAI o da parte di terzi soggetti; ii) per quelle concesse in licenza, una copia delle stesse fornita dal licenziante; iii) per quelle derivanti dalle interazioni degli utenti, l'acquisizione delle stesse durante tale fase e attraverso i feedback forniti dagli "addestratori" incaricati;

- di aver individuato nel legittimo interesse di cui all'art. 6, par. 1, lett. f), del Regolamento, la base giuridica per trattare i dati personali nelle fasi di addestramento e affinamento (fine-tuning) dei modelli e di aver svolto le opportune valutazioni sul livello di necessità del trattamento, in particolare per quanto riguarda i dati pubblicamente disponibili raccolti da Internet;

- per limitare ulteriormente il trattamento di dati personali, di fornire istruzioni ai collaboratori esterni, responsabili della categorizzazione dei dati, di escludere le informazioni che contengono dati personali dall'inclusione nel dataset di fine-tuning;

- di fornire agli utenti la possibilità di scegliere di non consentire l'utilizzo delle loro chat per i fini di addestramento del modello e di eliminare il proprio account, rimuovendo in modo permanente i dati associati.

b) riguardo alle informazioni sul trattamento dei dati fornite agli interessati:

- di aver predisposto, in aggiunta alle informazioni rese disponibili mediante l'apposita Notice, direttamente collegata alla privacy policy, uno specifico strumento per la rimozione dei dati personali, disponibile nel sito web della Società a vantaggio di chiunque voglia opporsi all'utilizzo dei propri dati per finalità di addestramento dei modelli;

- di aver pianificato il lancio di una campagna informativa sia online, che su due dei maggiori quotidiani italiani.

c) in merito al data breach:

- di aver pubblicato nel proprio sito web un post relativo all'evento occorso il 20 marzo 2023, fornendo informazioni sull'accaduto e di aver contattato direttamente via e-mail tutti gli interessati potenzialmente coinvolti dall'evento, compresi 440 interessati italiani;

- che il data breach avrebbe coinvolto alcuni dati personali di utenti sottoscrittori del servizio ChatGPT Plus ed in particolare nome, cognome, indirizzo e-mail e informazioni di fatturazione, tra cui la tipologia di carta di credito, le ultime quattro cifre della stessa e la relativa data di validità;

- di aver individuato il bug, sconosciuto in precedenza, in una libreria open source e di aver risolto il problema lo stesso giorno in cui è stato rilevato (20 marzo 2023).

Rispetto alla richiesta di informazioni supplementari dell'Autorità del 6 ottobre 2023, invece, la Società, con nota del 20 novembre 2023 (prot. n. 56039/23) ha rappresentato:

a) di intendere per fonte accessibile al pubblico ogni informazione gratuitamente ed apertamente disponibile in Internet e che, prima di utilizzare tali dati, applica dei filtri per rimuovere le informazioni da cui il modello non deve imparare, come hate speech, contenuti per adulti, siti di aggregazione di contenuti e spam;

b) che oltre il 99% dei dati di pre-training provengono da fonti pubbliche come Common Crawl, una delle più ampie fonti di parole liberamente accessibili nel web, mentre il rimanente 1% proviene da dati concessi in licenza da terze parti che posseggono dataset di alta qualità o che sono generati internamente dalla Società stessa; nella scelta di dataset sotto licenza, OpenAI seleziona fonti idonee a costruire modelli di intelligenza artificiale sicuri, come dataset di alta qualità che rispecchino diverse posizioni informative e che includano materie specialistiche come la scienza e la matematica;

c) che non vi è alcun coinvolgimento di Common Crawl, né delle terze parti che forniscono dataset, nella preparazione dei dataset stessi per il processo di addestramento;

d) di non poter escludere a priori informazioni personali dal training, perché necessarie per addestrare il modello su come funziona il linguaggio e per comprendere e rispondere alle domande (prompt) degli utenti;

e) che nessun dato personale viene usato a fini di profilazione, di pubblicità, di proposte di vendita o di informazione;

f) di aver implementato diverse misure di salvaguardia per ridurre il trattamento di dati personali nella fase di pre-training (prima fase in cui si insegna al modello la capacità di predire, ragionare e risolvere problemi), di post-training (seconda fase in cui si allinea il modello ai valori ed alle preferenze umane) e nelle risposte alle domande degli utenti (output), specificando quali operazioni vengono effettuate in concreto per ogni fase. OpenAI ha riferito di aver adottato procedure per una rivalutazione periodica dei processi di addestramento al fine di applicare al meglio il principio di minimizzazione, con specifico riferimento al filtraggio post-training [OMISSIS]. A tale riguardo OpenAI ha ribadito che offre agli utenti la possibilità di non concedere le proprie conversazioni per l'addestramento del modello e che li avvisa di non inserire informazioni sensibili nelle domande (prompt). La Società ha altresì riferito di aver implementato un web crawler (CPTBot) che esegue il crawling delle pagine web pubblicamente disponibili per cercare informazioni che possano

migliorare i modelli (i proprietari dei siti web possono disabilitare GPTbot inserendolo nel file robot.txt). La Società ha sottolineato di essere stata la prima AI company ad aver creato ed implementato tale approccio ed ha riferito che anche Common Crawl, che è la fonte primaria di GPT 3.5 e GPT 4, ha previsto degli strumenti a favore dei siti web per escludere il crawling da parte dei suoi bot. Per quanto riguarda le richieste di cancellazione degli utenti, OpenAI ha fatto riferimento ai form presenti online, alla possibilità offerta agli utenti di escludere le conversazioni dall'addestramento e di cancellare i loro account;

g) rispetto alla richiesta di specificare se sono stati implementati meccanismi quali entity detection, word embeddings (o simili) per identificare i dati personali pertinenti (sotto forma di nomi, posizioni, date) all'interno di flussi testuali, e come questi vengono elaborati durante la fase di training dell'algoritmo e durante la generazione di flussi testuali, OpenAI ha rinviato alla risposta fornita al punto f) di cui sopra;

h) quanto alle soluzioni tecniche implementate per consentire alle persone fisiche di esercitare il diritto di rettifica dei dati inesatti ed il diritto di cancellazione dei dati già utilizzati, OpenAI, fatta una premessa sul funzionamento di un LLM, ha fornito un riscontro descrittivo della gestione delle richieste di rettifica e cancellazione. Quanto alla rettifica, la Società ha riferito di correggere, ove possibile, le inesattezze (ad es. con il cd. fine-tuning) e che, se tale approccio non è sufficiente, data la complessità tecnica del modello, esclude tout court le informazioni personali dall'output. OpenAI ha precisato che i modelli GPT sono stati istruiti per rifiutare i prompt relativi ad informazioni private e sensibili, salvo si tratti di personaggi pubblici e, per illustrare tale funzione, ha fornito alcuni screenshot a titolo esemplificativo. Quanto alla cancellazione, OpenAI ha rinviato al form online descrivendo brevemente la procedura di cancellazione.

Unitamente al riscontro del 20 novembre 2023, la Società ha prodotto sei documenti, tra cui una copia della valutazione del legittimo interesse (Legitimate Interest Assessment - LIA) priva di data.

Nella medesima nota del 20 novembre 2023 (prot. n. 56039/23) OpenAI, con specifico riferimento alla verifica dell'età degli utenti, ha prodotto copia della valutazione di impatto (DPIA) redatta sul sistema di age verification adottato per i soli utenti in Italia, il contratto con Yoti, un whitepaper sull'age verification, la privacy policy aggiornata al 23 giugno 2023 e l'age verification notice.

La Società ha precisato altresì che, all'esito di un primo breve periodo di prova, ha ritenuto opportuno rettificare alcune delle soluzioni già prospettate, fornendo i seguenti dettagli di funzionamento:

1. Age estimation: la verifica dell'età si basa sulla stima della stessa effettuata attraverso un austoscatto (selfie) fornita dall'utente attraverso la App Yoti o il relativo sito. Il tool si limita a stimare l'età e la foto viene cancellata immediatamente;

2. ID scan: la verifica dell'età si basa sulla scansione di un documento di identità dell'utente (passaporto, patente di guida o carta di identità) e sul calcolo dell'età ricavata dalla data di nascita. Nell'ambito di tale verifica Yoti cancella il dato una volta completata la sessione o dopo 25 ore (tale ultima ipotesi si riferisce alle ipotesi di sessioni non finalizzate dall'utente), a seconda dell'evento che si verifica per primo;

3. Credit card: la verifica dell'età è legata all'utilizzo di una carta di credito intestata all'utente attraverso una transazione temporanea minima (£ 0.30) in relazione alla quale Yoti si limita a richiedere l'autorizzazione al pagamento.

Tutte le soluzioni sopra prospettate prevedono che Yoti comunichi ad OpenAI solo un ID di sessione e il risultato della verifica dell'età dell'utente.

Il trattamento di dati personali effettuato da Yoti, società statunitense qualificata come responsabile del trattamento, è regolato da un accordo ai sensi dell'art. 28 del GDPR, che è stato prodotto (i dati vengono trattati nel datacenter di Yoti nel Regno Unito).

OpenAI ha specificato che Yoti è conforme agli standard SOC 2 Tipo 2 (System and Organization Controls), ISO/IEC 27001:2013 (sistema di gestione della sicurezza delle informazioni), ISO 9001 (Sistema di gestione della qualità) e ISO/IEC 27701 (Sistema di Gestione della Conformità al GDPR), è stata certificata dalla Age Check Certification Service Ltd, un ente di certificazione terzo inglese ed è stata approvata dalla German Association for Voluntary Self-Regulation of Digital Media e dalla German Commission for Youth Media Protection.

3.1 ATTO DI COMUNICAZIONE DI AVVIO DEL PROCEDIMENTO AI SENSI DELL'ART. 166, CO. 5, CODICE E DELL'ART. 12, REGOLAMENTO DEL GARANTE N. 1/2019 E MEMORIE DIFENSIVE DELLA SOCIETÀ

Con nota del 26 gennaio 2024 (prot. 10531/24) l'Autorità ha notificato alla Società l'atto di comunicazione di avvio del procedimento per l'adozione dei provvedimenti correttivi e sanzionatori ai sensi dell'art. 166, comma 5, del Codice e dell'art.12 del Regolamento interno del Garante n.1/2019, contestando ad OpenAI la presunta violazione degli artt. 5; 6; 8; 12; 13; 24; 25 par. 1; 33; 83, par. 5, lett. d), del Regolamento in relazione al trattamento dei dati personali svolto dalla Società, attraverso il servizio ChatGPT, alla data del 30 marzo 2023.

Con nota dell'11 marzo 2024 (prot. 30480/24) OpenAI ha fornito riscontro alla nota di comunicazione di avvio del procedimento ed ha chiesto di essere audita ex art. 166, comma 6, del Codice e art. 13 del Regolamento del Garante n. 1/2019.

Di seguito si riportano le violazioni contestate e le relative controdeduzioni difensive della Società.

3.1.1 CON RIFERIMENTO AL DATA BREACH

A norma dell'art. 33 del Regolamento, la violazione dei dati personali deve essere notificata dal titolare all'autorità competente ai sensi dell'art. 55, del Regolamento senza ingiustificato ritardo e, comunque, entro 72 ore dall'evento, salvo che la stessa violazione non presenti un rischio per i diritti e le libertà delle persone fisiche. Qualora la violazione presenti un rischio elevato, il successivo art. 34 del Regolamento impone anche la notifica, senza ingiustificato ritardo, agli interessati.

Nel caso di specie, il Garante ha rilevato che non è pervenuta all'autorità alcuna notifica del data breach occorso il 20 marzo 2023, nonostante il coinvolgimento di interessati italiani e la sussistenza della propria competenza ai sensi dell'art. 55 del Regolamento, considerata l'inapplicabilità degli artt. 56 e ss. del Regolamento non esistendo, all'epoca dei fatti, uno stabilimento unico o principale della Società nell'Unione europea. Si aggiunga che dalla comunicazione effettuata dalla Società, a norma dell'art. 34 del Regolamento, agli interessati coinvolti nel data breach è possibile dedurre una ammissione della Società stessa in merito alla valutazione del rischio elevato legato all'evento.

Pertanto, nel caso di specie, l'Autorità, ha rilevato che OpenAI ha omesso di notificare al Garante l'evento di violazione dei dati occorso il 20 marzo 2023, ed ha contestato la presunta violazione dell'art. 33 del Regolamento.

A tale riguardo, nella memoria difensiva, OpenAI ha, tra l'altro, rappresentato:

- [OMISSIS]

- [OMISSIS]

- [OMISSIS]

- [OMISSIS]

- [OMISSIS]

- di aver descritto nel dettaglio, nella notifica dell'incidente trasmessa all'autorità di controllo irlandese (di seguito anche "IDPC"), il potenziale impatto della violazione su tutti gli Stati membri dell'UE, compresa l'Italia e che, sebbene l'incidente non comportasse un rischio elevato, di aver volontariamente intrapreso l'ulteriore iniziativa di pubblicare un post sul proprio sito web dal titolo "March 20 ChatGPT outage: Here's what happened" e di aver informato direttamente tutti gli utenti potenzialmente coinvolti in data 24 marzo 2023;

- di aver notificato la violazione alla IDPC, ritenendo che quest'ultima avrebbe condiviso le informazioni ricevute con gli altri membri del Comitato europeo dei Garanti (di seguito anche "EDPB" o "Comitato"), tra cui il Garante;

- di essere venuta a conoscenza dell'incidente di sicurezza mentre era in procinto di costituire la sua filiale irlandese, OpenAI Ireland Ltd., come suo unico stabilimento nell'Unione Europea;

- che prima che si verificasse l'incidente OpenAI stava lavorando all'apertura della filiale irlandese "OpenAI Ireland Ltd." e che quest'ultima è stata formalmente costituita pochi giorni dopo che OpenAI è venuta a conoscenza dell'incidente di sicurezza. Alla luce di ciò, la Società ha dichiarato di aver notificato l'incidente all'IDPC il 23 marzo 2023, entro 72 ore dal momento in cui ne è venuta a conoscenza, comunicando le informazioni relative al numero di soggetti potenzialmente coinvolti nell'Unione europea, nonché la ripartizione per Paese dell'UE dei soggetti potenzialmente coinvolti, inclusa l'Italia. Nel farlo, OpenAI ha dichiarato di aver presunto in buona fede che il Garante sarebbe venuto a conoscenza dell'incidente che avrebbe avuto tutte le informazioni pertinenti a sua disposizione;

- che il 13 aprile 2023 OpenAI ha ricevuto dalla IDPC una richiesta di informazioni in merito all'incidente, alla quale ha prontamente risposto il 27 aprile 2023;

- che, la Società ha valutato che l'incidente non comportasse l'applicazione dell'art. 34 del Regolamento.

3.1.2 CON RIFERIMENTO ALLA MANCATA INDIVIDUAZIONE DELLA CONDIZIONE DI LICEITÀ DEL TRATTAMENTO

L'art. 6 del Regolamento prescrive le condizioni di liceità del trattamento elencando le possibili basi giuridiche su cui il titolare deve fare affidamento per poter trattare lecitamente i dati personali necessari per lo svolgimento della propria attività. La base giuridica, come chiarito dall'EDPB (cfr. le Linea guida 2/2019 sul trattamento di dati personali ai sensi dell'art. 6, par. 1, lettera b), del regolamento generale sulla protezione dei dati nel contesto della fornitura di servizi online agli interessati) deve essere individuata prima dell'attuazione del trattamento e deve essere specificata nelle informazioni fornite agli interessati conformemente agli articoli 13 e 14.

Nel caso di specie, il Garante ha rilevato che il trattamento dei dati personali da parte di OpenAI per le finalità di addestramento del modello linguistico GPT sotteso al funzionamento del servizio ChatGPT è avvenuto ben prima che lo stesso servizio fosse reso disponibile al pubblico in data 30 novembre 2022 (cfr. https://openai.com/blog/chatgpt). Quanto meno a tale data, ma ragionevolmente anche in un periodo precedente, dunque, OpenAI avrebbe dovuto individuare una base giuridica su cui fondare il trattamento per l'addestramento del modello. Dalla documentazione in atti, invece, non emerge alcun elemento da cui far discendere la prova che la

Società abbia individuato una base giuridica in un momento anteriore all'inizio del trattamento, in particolare al momento della raccolta dei dati da Internet per l'addestramento del LLM di ChatGPT. Tale assunto risulta rafforzato anche dall'assenza di qualsivoglia indicazione, in merito alla base giuridica, nelle informazioni fornite agli interessati.

Lo stesso confuso riferimento al legittimo interesse o all'esecuzione contrattuale riportato al paragrafo 2.2 della nota della Società del 6 aprile 2023, appare indicativo di una inadeguata riflessione sul tema.

Il Garante ha osservato che OpenAI non ha fornito neanche elementi tali da escludere l'esigenza del consenso (una delle due opzioni indicate dal Garante al punto 5 del provvedimento n. 114/2023) quale base giuridica del trattamento, indicando invece sovente, nei riscontri forniti dalla Società, la possibilità di ricorrere a forme di opt-out.

Non sarebbero dirimenti, a parere dell'Autorità, a tal proposito, i riscontri forniti da OpenAI e, in particolare, la DPIA e la LIA prodotte rispettivamente in data 19 maggio e 20 novembre 2023:

a) la DPIA, che risulta originariamente redatta in una prima versione provvisoria ("Original draft date: February 24, 2023") in data 24 febbraio 2023 ed aggiornata in data 19 maggio 2023, fa riferimento al legittimo interesse come base giuridica del trattamento de quo, ma non fornisce alcun elemento utile a dimostrare che la relativa valutazione di adeguatezza abbia avuto luogo prima del 30 novembre 2022, in adempimento agli obblighi del titolare di cui all'art. 6 del Regolamento;

b) la LIA, a sua volta, è stata prodotta solo a fronte di espressa richiesta dell'Autorità (cfr. richiesta integrativa in data 6 ottobre 2023). Si tratta di un documento non datato che non fornisce alcun elemento che consenta di ritenere che l'individuazione di una delle basi giuridiche indicate nell'art. 6 del Regolamento sia avvenuta in data precedente al 30 marzo 2023.

Pertanto, nel caso di specie, atteso che OpenAI non è stata in grado di dimostrare e comprovare ai sensi dell'art. 5, par. 2, del Regolamento (principio di responsabilizzazione), di aver chiaramente individuato sino alla data del 30 marzo 2023 e, in ogni caso, prima dell'inizio dell'attività di trattamento, una base giuridica per il trattamento dei dati personali per finalità di addestramento del modello GPT sotteso al funzionamento del servizio ChatGPT, reso disponibile al pubblico già a far data dal 30 novembre 2022, l'Autorità ha contestato la presunta violazione degli artt. 5, par. 2 e 6 del Regolamento.

A tale riguardo, nella memoria difensiva, il titolare ha, tra l'altro, rappresentato che:

- OpenAI ha rilasciato ChatGPT il 30 novembre 2022 come anteprima di ricerca gratuita, al fine di far progredire la ricerca e lo sviluppo dell'IA, promuovere un uso responsabile dell'AI e raccogliere feedback da un insieme diversificato di utenti. OpenAI non aveva pianificato o previsto che diventasse un importante servizio rivolto ai consumatori, invece il pubblico ha scoperto ChatGPT e ha risposto in modo positivo, utilizzandolo per una vasta gamma di attività benefiche. Una volta che ChatGPT è diventato un servizio disponibile per soggetti localizzati nell'UE, OpenAI ha acquisito nuovi obblighi ai sensi del GDPR, tra cui quello di documentare la valutazione degli interessi legittimi ("LIA");

- prima del 30 novembre 2022 OpenAI non aveva sede nell'UE, non offriva beni o servizi agli interessati localizzati nell'UE e non monitorava il loro comportamento. All'epoca, il trattamento dei dati personali da parte di OpenAI era in linea con i principi di protezione dei dati, ma non era soggetto al GDPR. In qualità di leader nella comunità della ricerca scientifica per il suo lavoro di avanzamento della ricerca sull'IA, OpenAI è stata a lungo

trasparente sul trattamento dei dati disponibili al pubblico per addestrare i modelli di IA, anche in documenti tecnici, schede dei modelli e altre pubblicazioni, in cui ha anche descritto le sue valutazioni e le azioni di mitigazione dei rischi adottate;

- quando OpenAI ha lanciato ChatGPT come anteprima di ricerca gratuita il 30 novembre 2022, si trattava di un'interfaccia che permetteva alle persone di interagire con la ricerca di OpenAI. Una volta che gli interessati localizzati nell'UE lo hanno scoperto e utilizzato come servizio, il GDPR, compreso l'obbligo di identificare una base giuridica e completare una LIA, è diventato applicabile al trattamento dei dati da parte di OpenAI;

- l'11 gennaio 2023, OpenAI e i suoi consulenti legali esterni hanno iniziato a lavorare sulla documentazione della LIA come parte della valutazione d'impatto sulla protezione dei dati ("DPIA") e una prima bozza è stata completata il 24 febbraio 2023. Questa bozza documentava la base giuridica di OpenAI, cioè il suo interesse legittimo, per il trattamento dei dati delle conversazioni con ChatGPT a fini di addestramento comprese le misure di salvaguardia che OpenAI aveva implementato;

- il 14 marzo 2023, OpenAI ha aggiornato la sua Informativa Privacy per menzionare esplicitamente che OpenAI si basa sui suoi legittimi interessi (...) quando sviluppiamo (...) i nostri Servizi;

- il riferimento alla base contrattuale come condizione di liceità riguarda il trattamento dei dati degli utenti per fornire il servizio ChatGPT;

- con riferimento alla base giuridica del consenso, il punto 5 del Provvedimento n. 114 non ha richiesto a OpenAI di dimostrare perché il consenso non sia la base giuridica più appropriata per l'addestramento dei modelli di IA, altrimenti in tale occasione la Società avrebbe approfondito le difficoltà pratiche nell'ottenere il consenso nel contesto dell'addestramento di modelli di IA e fatto riferimento agli orientamenti normativi pertinenti che sostengono che l'addestramento di modelli di IA possa basarsi sulla base giuridica dell'interesse legittimo.

3.1.3 CON RIFERIMENTO ALLA PRIVACY POLICY (VERSIONE AGGIORNATA AL 14 MARZO 2023)

L'analisi svolta dal Garante si è limitata alle attività di trattamento e agli adempimenti posti in essere da OpenAI alla data del 30 marzo 2023 e, di conseguenza, la privacy policy valutata è quella in vigore a tale data, ovverosia la versione aggiornata al 14 marzo 2023.

Il funzionamento del servizio ChatGPT implica due tipologie di trattamenti: una limitata ai dati degli utenti del servizio, dati richiesti per l'iscrizione al servizio e utilizzati nell'interazione con la piattaforma; una seconda costituita dall'addestramento del modello GPT – posto alla base del servizio offerto – e che coinvolge i dati disponibili in rete riferibili anche ad interessati terzi non utilizzatori del servizio. Ne deriva che tali trattamenti richiedono che il titolare fornisca idonee informazioni sia ai soggetti che si iscrivono ed utilizzano il servizio (utenti), sia ai soggetti che non fruiscono del servizio ed i cui dati sono trattati per permetterne il funzionamento (non utenti).

Alla data del 30 marzo 2023 il titolare ha reso disponibile la privacy policy solo in lingua inglese (anche per i minori) e non facilmente reperibile nel sito web della Società in quanto, nel flusso di registrazione, non ha pubblicato il link alla informativa in una posizione tale da permetterne la lettura prima dell'inserimento dei dati per la creazione di un account utente.

Dal punto di vista contenutistico, alla data del 30 marzo 2023, OpenAI ha fornito informazioni sul trattamento esclusivamente con riferimento ai dati personali che la Società raccoglie dagli utenti quando questi utilizzano i servizi offerti (cfr. preambolo alla privacy policy nella versione

aggiornata al 14 marzo 2023). Non ha, invece, fornito nessuna informazione in relazione al trattamento dei dati personali dei non utenti, che vengono trattati per l'addestramento del LLM. Per tali carenze, l'Autorità ha contestato la presunta violazione degli artt. 5, par. 1, lett. a), 12 e 13 del Regolamento.

Il Garante ha altresì rilevato che nella privacy policy al 14 marzo 2023 non è presente alcuna informazione in merito alle operazioni di trattamento relative ai dati degli utenti durante l'utilizzo del servizio e finalizzate all'addestramento del modello GPT. Infatti, le informazioni contenute nel par. 1 della citata privacy policy relative agli Usage Data, nonché il riferimento di cui al par. 2 ad una generica finalità di fornitura e miglioramento dei servizi non appaiono idonee ad informare adeguatamente gli utenti del servizio che l'addestramento dell'algoritmo avviene anche sulla base dal trattamento dei dati che gli stessi condividono semplicemente utilizzando il servizio ChatGPT.

Nello specifico, la prima parte del par. 2 della privacy policy fornisce le seguenti informazioni:

> 2. How we use personal information. We may use Personal Information for the following purposes:
>
> • To provide, administer, maintain, improve and/or analyze the Services;
>
> • To conduct research;
>
> • To communicate with you;
>
> • To develop new programs and services;
>
> • To prevent fraud, criminal activity, or misuses of our Services, and to ensure the security of our IT systems, architecture, and networks; and
>
> • To comply with legal obligations and legal process and to protect our rights, privacy, safety, or property, and/or that of our affiliates, you, or other third parties […].

Deve rilevarsi che la formulazione di tali informazioni risulta poco chiara, in quanto consiste in un elenco di finalità del trattamento riferite, senza distinzioni, a tutti i dati personali raccolti, nonostante tali dati - elencati al precedente par. 1 - si dividano in due categorie: dati forniti dagli utenti (Account Information, User Content, Communication Information, Social Media Information) e dati raccolti automaticamente dal servizio (Log Data, Usage Data, Device Information, Cookies e Analytics).

Tale impostazione non appare corretta. A titolo esemplificativo, per il perseguimento della finalità to communicate with you è evidentemente necessario il trattamento solo di alcune delle informazioni (es. nome e account, dati di contatto) che vengono riportate nella informativa proposta agli interessati. Inoltre, per ogni diversa finalità di trattamento, avrebbe dovuto essere riportata la corrispondente tipologia di dati, non essendo possibile richiamare tutti i dati personali raccolti e limitarsi ad elencare tutte le finalità. Di contro, l'impostazione adottata dal titolare non è idonea a chiarire agli interessati le specifiche finalità per cui è trattata ciascuna categoria di dati raccolta. Tra le finalità indicate, OpenAI inserisce altresì un generico e non meglio specificato to conduct research che non fornisce alcuna precisa informazione circa la natura e il perimetro di tale finalità.

Nella parte finale del par. 2 la Società riferisce anche di trattare informazioni de-indentificate in forma anonima o de-identificata; al di là della tautologia della definizione proposta, la questione attiene, ancora una volta, a profili di chiarezza e di correttezza - anche tecnica - delle informazioni fornite. Infatti, un'informazione anonima, di per sé, non costituisce dato personale in quanto non permette la re-identificazione dell'interessato (si parla, infatti, di trattamento irreversibile).

L’informazione fornita agli interessati nella privacy policy non risulta pertanto corretta.

Parimenti risulta di scarsa chiarezza l’informazione di cui al par. 8 in cui si parla dell’anonimizzazione o de-identificazione delle informazioni personali so that it can no longer be associated with you. Come già evidenziato, laddove le informazioni siano anonime non è possibile risalire in nessun caso all’interessato originario; di contro, laddove le informazioni siano “de-identificate” è possibile, tramite un processo inverso, la re-identificazione dell’interessato.

Quanto sopra rappresentato ha portato l’Autorità a contestare la presunta violazione degli artt. 5, par. 1, lett. a), 12 e 13 del Regolamento anche nel contesto dei trattamenti dei dati degli utenti del servizio.

Nella memoria difensiva il titolare ha, tra l’altro, rappresentato che OpenAI è trasparente in relazione al trattamento dei dati personali e ha informato gli interessati non utenti sin dal 2016 mediante la pubblicazione di oltre 170 diversi post, documenti di ricerca, articoli e comunicazioni al fine di istruire le persone su vari argomenti relativi all’IA ed allo sviluppo dei propri modelli, facilmente accessibili tramite un indice di ricerca.

Con riferimento alle informazioni rese agli utenti la Società ha, tra l’altro, chiarito:

- di aver informato gli utenti in merito al trattamento delle loro conversazioni per la calibrazione (fine-tuning) dei propri modelli GPT. In particolare, nella privacy policy aggiornata al 14 marzo 2023 il trattamento di tali dati è stato descritto sotto la voce Quali dati personali raccogliamo, come segue: Contenuti dell’utente: Quando l’utente utilizza i nostri Servizi, raccogliamo i Dati personali inclusi negli input, nei file caricati o nei feedback che lutente fornisce ai nostri Servizi ("Contenuti"). Nella stessa informativa le finalità di utilizzo dei Contenuti dell’Utente da parte di OpenAI sono state chiaramente descritte alla voce Fornire, amministrare, mantenere, migliorare e/o sviluppare i Servizi e Condurre ricerche, relative all’utilizzo delle conversazioni di ChatGPT per la messa a punto dei modelli alla base di ChatGPT;

- che sin dal lancio di ChatGPT, il 30 novembre 2022, contattando i canali di assistenza di OpenAI gli utenti hanno potuto scegliere di rinunciare all’utilizzo delle loro conversazioni per il fine-tuning dei modelli di OpenAI e che, a partire da febbraio 2023, OpenAI ha semplificato la procedura mettendo a disposizione degli utenti un modulo online per la rinuncia all’utilizzo delle conversazioni per tale finalità;

- che sin dal 30 novembre 2022, l’interfaccia di ChatGPT ha informato gli utenti con un pop-up ben visibile (prima che gli stessi interagiscano con ChatGPT) che Le conversazioni possono essere esaminate dai nostri addestratori di IA per migliorare i nostri sistemi e di non condividere alcuna informazione sensibile nelle conversazioni;

- che la finalità perseguita da OpenAI è la ricerca nel campo dell’IA (le capacità dei modelli, la sicurezza e il loro allineamento, il linguaggio, il ragionamento, l’IA responsabile, l’apprendimento con feedback umano, l’interpretabilità e altri argomenti);

- che, considerata l’elevata soglia normativa necessaria per poter parlare di anonimizzazione, il riferimento alle informazioni aggregate e de-identificate è inteso dalla Società come ad informazioni non completamente anonimizzate, quindi di carattere personale.

In relazione ai non utenti la Società ha, tra l’altro, chiarito che:

- sebbene il GDPR non fosse applicabile [prima del 30 novembre 2023, n.d.r.] alla raccolta dei dati in questione, OpenAI è sempre stata trasparente in merito al trattamento dei dati

disponibili pubblicamente per addestrare i propri modelli. Ad esempio, quando ha annunciato il GPT-2 (annunciato per la prima volta nel febbraio 2019), OpenAI ha pubblicato un post, un documento tecnico e la scheda del modello GPT-2 che illustravano in dettaglio l'uso da parte di OpenAI di dati Internet pubblici per l'addestramento dei suoi modelli linguistici. Quando OpenAI ha annunciato il GPT-3 nel giugno 2020, ha descritto in dettaglio l'uso di fonti di dati pubblici per addestrare i suoi modelli (si veda il documento tecnico e la scheda del modello GPT-3). Anni prima di rilasciare ChatGPT come anteprima di ricerca gratuita nel novembre 2022, OpenAI era già leader nella comunità scientifica per il suo lavoro di avanzamento della ricerca sull'IA, anche per quanto riguarda la pubblicazione delle sue scoperte sull'addestramento di modelli linguistici di grandi dimensioni con fonti di dati pubbliche. Quando OpenAI ha annunciato il GPT-4 nel marzo 2023, ha pubblicato anche un post, un rapporto tecnico e la scheda di sistema del GPT-4. Coerentemente con il suo ruolo di organizzazione di ricerca, OpenAI ha pubblicato queste informazioni in modo chiaro e visibile nelle pubblicazioni di ricerca, anche prima di pubblicare l'Informativa "How ChatGPT and our language models are developed", fornita con collegamento ipertestuale diretto all'inizio dell'informativa Privacy a partire da aprile 2023.

3.1.4 CON RIFERIMENTO ALL'ASSENZA DI MECCANISMI DI VERIFICA DELL'ETÀ DEI MINORI

L'art. 24, par. 1, della Carta dei diritti fondamentali dell'uomo stabilisce che i minori hanno diritto alla protezione ed alle cure necessarie per il loro benessere. Il paragrafo 2 della stessa norma afferma che in tutti gli atti compiuti da autorità pubbliche o da istituzioni private, l'interesse del minore deve essere preso in considerazione in via prioritaria.

Il medesimo principio è racchiuso nel considerando 38 del Regolamento che recita: i minori meritano una specifica protezione relativamente ai loro dati personali, in quanto possono essere meno consapevoli dei rischi, delle conseguenze e delle misure di salvaguardia interessate nonché dei loro diritti in relazione al trattamento dei dati personali.

La normativa europea in materia di protezione dei dati personali richiede al titolare del trattamento di adottare adeguate misure tecniche e organizzative volte ad attuare in modo efficace i principi di protezione dei dati e di integrare nel trattamento le garanzie necessarie al fine di soddisfare i requisiti previsti dal Regolamento e tutelare i diritti degli interessati.

Ai sensi dell'art. 24, par. 1, del Regolamento, Tenuto conto della natura, dell'ambito di applicazione, del contesto e delle finalità del trattamento, nonché dei rischi aventi probabilità e gravità diverse per i diritti e le libertà delle persone fisiche, il titolare del trattamento mette in atto misure tecniche e organizzative adeguate per garantire, ed essere in grado di dimostrare, che il trattamento è effettuato conformemente al presente regolamento. Dette misure sono riesaminate e aggiornate qualora necessario.

Ai sensi dell'art. 25 del Regolamento, il titolare deve adottare tali misure tenendo conto dello stato dell'arte e dei costi di attuazione, nonché della natura, dell'ambito di applicazione, del contesto e delle finalità del trattamento, come anche dei rischi aventi probabilità e gravità diverse per i diritti e le libertà delle persone fisiche costituiti dal trattamento, sia al momento di determinare i mezzi del trattamento sia all'atto del trattamento stesso.

Con le linee guida n. 4/2019 sull'art. 25 del Regolamento l'EDPB ha chiarito che il fulcro della disposizione è garantire una adeguata ed efficace protezione dei dati fin dalla progettazione e una protezione per impostazione predefinita, il che significa che i titolari dovrebbero essere in grado di dimostrare che incorporano nel trattamento le misure e le garanzie adeguate ad assicurare l'efficacia dei principi di protezione dei dati, dei diritti e delle libertà degli interessati. E, nell'ambito delle raccomandazioni, il Comitato ha invitato i titolari a tenere conto, nell'ambito della

progettazione e impostazione del trattamento in un'ottica privacy oriented, anche degli obblighi di fornire una tutela specifica ai minori e ad altri gruppi vulnerabili.

Nel corso dell'istruttoria l'Autorità ha accertato che la natura del servizio offerto da OpenAI è inquadrabile nella definizione fornita all'art. 4, par. 1, punto 25, del Regolamento tra i servizi della società dell'informazione, che le attività di trattamento coinvolgono dati di minori e che, pertanto, la Società è tenuta ad assicurare una specifica protezione in favore dei minori, tenendo conto che i diritti e le libertà di questi ultimi debbono essere considerati prevalenti rispetto ad altri interessi coinvolti nelle attività di trattamento. Ai sensi della normativa il titolare del trattamento deve adoperarsi in ogni modo ragionevole per verificare che il consenso del minore sia prestato o autorizzato dall'esercente la responsabilità genitoriale, in considerazione delle tecnologie disponibili (art. 8 del Regolamento).

L'Autorità ha rilevato che, sino al 30 marzo 2023, OpenAI non ha adottato alcuna misura tecnica ed organizzativa atta a garantire il rispetto dei principi generali del Regolamento e la protezione dei diritti e delle libertà dei minori, esponendo pertanto i soggetti minorenni (ivi compresi quelli che, stando alle condizioni di servizio, non possono accedere e fruire dello stesso servizio perché minori di 13 anni) a rischi significativi per la loro persona, tra cui risposte inidonee rispetto ai loro gradi di sviluppo psicofisico e di autoconsapevolezza, che la normativa in argomento mira a tutelare.

Come noto, sino al 30 marzo 2023, la Società non ha implementato alcun meccanismo per la verifica dell'età degli utenti all'atto dell'iscrizione al servizio ChatGPT, ancorché le condizioni di servizio (ToS), nella versione a quel tempo vigente e aggiornata al 14 marzo 2023, individuassero i minorenni tra i potenziali utenti, stabilendo che, ai fini del perfezionamento di un valido vincolo contrattuale, per gli utenti di età compresa tra i 13 ed i 18 anni è necessario il consenso del titolare della responsabilità genitoriale. A tutela dei minori, la privacy policy, nella versione pubblicata online in data 30 marzo 2023 ed aggiornata al 14 marzo 2023, prevede solo un sistema di segnalazione (tramite posta elettronica) di eventuali conferimenti di dati personali da parte di soggetti minori di 13 anni attraverso il servizio.

Dalla documentazione in atti risulta che la Società ha manifestato il proprio impegno a valutare l'implementazione di meccanismi di verifica dell'età solo a seguito della misura correttiva imposta dall'Autorità nell'ambito del provvedimento di sospensione della limitazione provvisoria n. 114 dell'11 aprile 2023. In particolare, nel corso delle interlocuzioni che hanno fatto seguito all'adozione del provvedimento del 30 marzo 2023, con specifico riferimento al profilo dell'age verification, l'Autorità ha chiesto alla Società di implementare, non oltre il 30 aprile 2023, un meccanismo per richiedere a tutti gli utenti che si collegano dall'Italia, ivi inclusi quelli già registrati, di superare, in sede di primo accesso, un age gate che escluda, sulla base dell'età dichiarata, gli utenti minorenni.

Prima dell'intervento dell'Autorità, quindi, tutti gli utenti, compresi i minorenni, potevano iscriversi al servizio ChatGPT ed utilizzarlo senza alcuna previa verifica dell'età. Atteso che l'assenza di uno standard comune idoneo a garantire, in maniera certa e assoluta, l'efficacia di un modello di verifica dell'età dell'utente e la discussione tutt'ora in atto a livello europeo al riguardo, non possono essere considerate cause idonee ad escludere l'adempimento degli obblighi a cui è tenuto il titolare del trattamento, tra i quali, in primis, quello di minimizzare, attraverso uno sforzo ragionevole e considerando le tecnologie disponibili a tal scopo, i rischi a cui sono esposti i minori nell'ambito delle attività di trattamento dei dati personali. In relazione a tale profilo, l'Autorità ha contestato la violazione degli artt. 8, 24 e 25, par. 1 del Regolamento per la mancata predisposizione di idonei sistemi atti a verificare l'età dei soggetti alla data del 30 marzo 2023.

Con riferimento ai meccanismi di age verification, nella memoria difensiva, il titolare ha, tra l'altro, rappresentato che:

- i servizi di OpenAI, incluso ChatGPT, non sono rivolti ai minori, come chiarito nei termini di servizio, nella privacy policy e nell'Help Center di OpenAI;

- OpenAI sviluppa i propri servizi nell'ottica della propria missione di creare un'IA artificiale sicura e vantaggiosa per tutta l'umanità;

- OpenAI non incoraggia o incentiva gli utenti a pubblicare o condividere contenuti online né tratta i dati per vendere servizi, proporre annunci o profilare le persone;

- OpenAI ha implementato misure per limitare la capacità di ChatGPT di generare output indesiderati per tutti gli utenti (tra cui contenuti odiosi, molesti, violenti o per adulti, tra le altre categorie);

- per utilizzare ChatGPT gli utenti devono avere almeno 13 anni;

- dal 30 novembre 2022 la Società ha invitato, tramite l'informativa privacy, gli utenti a segnalare ad un indirizzo di posta elettronica dedicato i casi in cui vi sia fondato motivo di ritenere che un minore di 13 anni abbia fornito informazioni personali attraverso ChatGPT;

- OpenAI ha implementato una serie di salvaguardie per proteggere tutti gli utenti di ChatGPT, compresi gli utenti minori di 18 anni;

- la legge italiana e il GDPR non richiedono l'implementazione di un meccanismo di age verification;

- con l'intento di andare oltre le prassi di mercato per applicare un elevato livello di protezione, OpenAI ha deciso di conformarsi volontariamente ai punti 7 e 8 del Provvedimento n. 114.

Con riferimento all'applicabilità dell'art. 8 del Regolamento, nella memoria difensiva, il titolare ha, tra l'altro, rappresentato che:

- l'art. 8 del Regolamento non si applica alle attività di trattamento di OpenAI, in quanto la norma trova applicazione solo con riferimento a trattamenti basati sul consenso (ossia ai trattamenti la cui base giuridica è riconducibile al disposto di cui all''art. 6, par. 1, lett. a), del Regolamento);

- per fornire e mantenere il servizio ChatGPT, OpenAI si è basata sulla base giuridica contrattuale di cui all'art. 6, par. 1, lettera b), del GDPR; per addestrare i suoi modelli, OpenAI si è basata sui suoi legittimi interessi e su quelli di terzi e della società in generale, per costruire un'IA sicura e vantaggiosa a beneficio di tutta l'umanità, ai sensi dell'articolo 6, par. 1, lettera f), del GDPR. OpenAI ha inoltre invocato la base giuridica del suo legittimo interesse laddove il trattamento dei dati personali degli utenti fosse necessario (i) per prevenire frodi, attività criminali o abusi dei suoi servizi e per proteggere la sicurezza dei suoi sistemi e servizi, o (ii) per proteggere i diritti, la privacy, la sicurezza o la proprietà di OpenAI o dei suoi utenti, affiliati o terzi.; per adempiere a un obbligo legale, come ad esempio conservare le informazioni sulle transazioni per adempiere ai propri obblighi di registrazione, OpenAI si è basata sull'articolo 6(1)(c) del GDPR.

Con riferimento alle misure adottate per proteggere gli utenti, compresi i minori, nella memoria difensiva, il titolare ha, tra l'altro, osservato che:

- i termini di servizio aggiornati al 14 marzo 2023 e disponibili online il 30 marzo 2023 prevedono che Qualora l'utente avesse meno di 18 anni, dovrà avere il permesso dei genitori o del tutore legale per utilizzare i Servizi. L'uso del termine "permesso" nella frase di

cui sopra non si riferisce al consenso ai sensi dell'articolo 6, par. 1, lettera a), del GDPR. Si tratta di garantire che un adolescente confermi di avere il permesso dei genitori o del tutore legale per utilizzare ChatGPT, anche per sostenere la validità del contratto tra OpenAI e l'adolescente;

- già prima del 30 marzo 2023 la Società ha adottato misure tecniche e organizzative a tutela di tutti gli utenti, compresi i minori e altri gruppi vulnerabili, tra cui: i) controlli sulla sicurezza dell'output, (ii) controlli sulla sicurezza dei minori, (iii) controlli sulla somiglianza tra l'output del modello e i contenuti generati dall'uomo, (iv) controlli sull'imprecisione dell'output del modello, (v) attenuazione del rischio di pregiudizi (cd. bias), (vi) trasparenza e (vii) controlli sui dati.

In particolare, nell'ambito dei controlli sulla sicurezza dell'output il titolare ha dichiarato che, come descritto nella scheda di sistema GPT-4 del 23 marzo 2023, nella fase di pre-addestramento, OpenAI ha lavorato per identificare e rimuovere le informazioni indesiderate dai dataset di addestramento, comprese alcune categorie di siti web che sono noti per contenere informazioni imprecise, inaffidabili o potenzialmente dannose, come siti contenenti contenuti pirata o altri contenuti illegali, discorsi di odio, contenuti per adulti, siti che aggregano principalmente informazioni personali e spam. E che durante la fase di post-addestramento, OpenAI ha testato e sottoposto ad un gruppo di esperti (cd. red-team) i suoi modelli per verificare la presenza di output potenzialmente dannosi, per rifiutare richieste di contenuti potenzialmente dannosi e a fornire risposte sicure. Inoltre la Società ha dichiarato di aver sviluppato una suite di classificatori basati su regole ed apprendimento automatico per identificare i contenuti problematici che potrebbero violare le policy di OpenAI e di aver implementato molteplici livelli di misure per limitare la capacità di ChatGPT di generare output indesiderati per tutti gli utenti, compresi gli utenti minori di 18 anni.

Nell'ambito dei controlli sulla sicurezza dei minori dell'output OpenAI ha dichiarato che, prima del 30 marzo 2023, ha preso provvedimenti per limitare e rimuovere i contenuti dannosi dal testo di addestramento, tra cui contenuti erotici inappropriati e contenuti illegali, come quelli pedopornografici (child sexual abuse material, CSAM). Inoltre, la Società ha dichiarato di aver implementato filtri in ingresso e in uscita, liste di blocco, sistemi di identificazione e riduzione di contenuti inappropriati e illegali nei dataset di pre-addestramento e addestramento di ChatGPT, affinché il sistema non fornisca risposte potenzialmente dannose. E che le Usage policies di OpenAI, come disponibili al 30 marzo 2023, vietavano l'uso dei suoi modelli per CSAM o per qualsiasi contenuto che sfruttasse o danneggiasse i minori, e chiarivano che tali contenuti sarebbero stati segnalati al NCMEC [National Center for Missing & Exploited Children, n.d.r.].

Con riferimento alle misure adottate per far fronte ai rischi derivanti dalla somiglianza tra l'output del modello ed i contenuti generati dall'uomo, OpenAI ha dichiarato di aver addestrato i modelli alla base di ChatGPT affinché ricordino agli utenti, nelle risposte fornite, che ChatGPT è un servizio di intelligenza artificiale che si basa su di un modello linguistico e di aver previsto nei termini di servizio e nelle policy di utilizzo, disponibili al 30 marzo 2023, che gli utenti non possono dichiarare che l'output è stato generato da un essere umano.

Con riferimento alle misure adottate per far fronte all'imprecisione dell'output del modello, la Società ha dichiarato di aver implementato misure tecniche e organizzative secondo un approccio basato sul rischio in conformità agli articoli 24 e 25 GDPR.

Con riferimento alle misure adottate per mitigare il rischio di pregiudizi (cd. bias) OpenAI ha dichiarato di aver adottato, prima del 30 marzo 2023, una serie di misure per contribuire a mitigare tale rischio, compresi lo sviluppo e l'impiego di tecniche per rendere più probabile che i modelli rifiutino di produrre contenuti non sicuri o distorti, in conformità con le sue politiche, l'adozione di misure per ridurre i contenuti inappropriati, discutibili o non rappresentativi che possono essere presenti nei dati di addestramento del modello, la verifica degli output del modello per potenziali

distorsioni e la messa a punto dei modelli per affrontare le potenziali fonti di distorsione.

Con riferimento alle misure adottate per garantire la trasparenza, la Società ha dichiarato che, prima del 30 marzo 2023, a tutti gli utenti sono state presentate informazioni chiare nell'interfaccia di ChatGPT attraverso una finestra pop-up, che informava gli utenti, con un linguaggio chiaro e comprensibile, circa le principali funzionalità di ChatGPT, ovvero che può intrattenere conversazioni, ma che ha anche dei limiti, in quanto può occasionalmente generare informazioni non corrette, può occasionalmente produrre istruzioni dannose o contenuti tendenziosi e che ha un limite di conoscenza. OpenAI ha altresì dichiarato che agli utenti è stato anche ricordato in una serie di finestre pop-up che ChatGPT era un'anteprima di ricerca e che, sebbene ChatGPT abbia delle salvaguardie in atto, il sistema può occasionalmente generare informazioni imprecise e che ChatGPT non è destinato a dare consigli. Inoltre, gli utenti sono stati informati del fatto che le loro conversazioni con ChatGPT potrebbero essere riviste o utilizzate per migliorare i sistemi di OpenAI.

Con riferimento al controllo sui dati, la Società ha dichiarato che oltre a presentare agli utenti, compresi quelli di età inferiore ai 18 anni, un chiaro avviso che le conversazioni possono essere riviste e utilizzate per il miglioramento del modello e che non devono inserire informazioni sensibili, OpenAI ha anche fornito agli utenti opzioni per controllare l'uso dei dati delle loro conversazioni. Gli utenti, compresi quelli di età inferiore ai 18 anni, hanno avuto la possibilità di opporsi al trattamento delle loro conversazioni ChatGPT ai fini di addestramento fin dall'inizio, quando ChatGPT è stato rilasciato per la prima volta il 30 novembre 2022. Inizialmente, gli utenti potevano rivolgersi all'assistenza di OpenAI per opporsi al trattamento. Con la crescita del numero di utenti di ChatGPT e l'espansione delle operazioni di OpenAI, nel febbraio 2023 OpenAI ha creato un modulo online (Allegato 4 della lettera del 4 aprile 2023) per consentire agli utenti di opporsi all'utilizzo dei loro contenuti per addestrare e migliorare i modelli di OpenAI. Il titolare ha infine dichiarato che Per gli utenti che non si sono opposti all'utilizzo dei loro dati per le finalità di addestramento, OpenAI ha adottato misure di protezione della privacy-by-design, come prescritto dall'articolo 25 GDPR, prima che le conversazioni potessero essere utilizzate nel processo di addestramento. In particolare, OpenAI ha dissociato le conversazioni dall'account dell'utente, ha utilizzato un processo di filtraggio per rimuovere gli identificatori personali che gli utenti potrebbero aver inserito nelle loro conversazioni e ha incorporato una fase di revisione umana per confermare che qualsiasi richiesta con informazioni identificabili o sensibili fosse esclusa dai dataset etichettati che possono essere utilizzati per l'addestramento del modello. I dati delle conversazioni possono essere utilizzati per la messa a punto per migliorare la funzionalità e la sicurezza del modello solo dopo le fasi di de-identificazione.

3.1.5 CON RIFERIMENTO ALLA "GENERAZIONE" DI DATI INESATTI

Nel provvedimento di limitazione provvisoria il Garante ha rilevato che gli output del servizio ChatGPT possono risultare inesatti, fornendo false rappresentazioni delle persone.

La Società, a seguito delle richieste del Garante contenute nel provvedimento n. 114/2023, pur evidenziando le difficoltà tecniche, ha approntato alcune parziali misure per ridurre tali effetti ed ha espresso la volontà di procedere verso il costante miglioramento dei risultati.

Atteso che la "generazione" di dati personali inesatti o non aggiornati si pone in contrasto con, il principio di esattezza (accuracy), il quale stabilisce che i dati trattati siano "esatti e, se necessario aggiornati" e che debbano essere "adottate tutte le misure ragionevoli per cancellare o rettificare tempestivamente i dati inesatti", l'Autorità ha contestato la presunta violazione dell'art. 5, par. 1, lett. d), del Regolamento.

Nella memoria difensiva il titolare ha, tra l'altro, rappresentato che:

- OpenAI fornisce agli utenti informazioni evidenti e visibili circa la possibilità che fornisca output imprecisi. Queste informazioni si trovano all'interno di ChatGPT, nei termini di utilizzo e nelle policy di OpenAI, nel blog che annuncia ChatGPT e negli articoli dell'Help Center.

- OpenAI si è impegnata a migliorare l'accuratezza dei suoi modelli ed è stata trasparente riguardo ai processi di addestramento, test e validazione dei modelli di OpenAI per quanto riguarda l'accuratezza. Ad esempio, OpenAI ha pubblicato informazioni che spiegano come essa abbia valutato la potenziale imprecisione del suo modello GPT-4 in varie pubblicazioni di ricerca, come la scheda di sistema GPT-4 e la relazione tecnica GPT-4.

- vi sono studi, che sono stati citati, che dimostrano che i modelli GPT 3.5 e GPT 4 sottesi al servizio ChatGPT hanno un tasso di allucinazione più basso rispetto ai principali competitor di servizi di intelligenza artificiale generativa.

La Società ha dichiarato che sin dal suo lancio nel novembre 2022 ChatGPT mostra agli utenti prima dell'utilizzo del servizio un pop-up contenente informazioni sulla potenziale imprecisione degli output di ChatGPT; la Società ha precisato che le stesse informazioni sono reperibili, sin dal novembre 2022, nelle FAQ del sito web, in un post sul blog di presentazione del servizio e dal marzo del 2023 sono pubblicate nei termini di servizio.

OpenAI ha altresì riferito che gli utenti hanno la possibilità di segnalare le inesattezze e richiederne la rettifica. In particolare, l'interessato può chiedere a OpenAI di correggere un'imprecisione che lo riguarda scrivendo ad una cesella di posta elettronica dedicata. A seguito di tale segnalazione OpenAI tenta di correggere l'imprecisione (ad esempio, affinando il modello), oppure, se non è in grado di farlo a causa della complessità tecnica del funzionamento dei suoi modelli, impedisce che le informazioni personali della persona appaiano nell'output di ChatGPT. Tale procedimento è descritto privacy policy di OpenAI.

Infine, la Società ha ribadito che sta investendo molto per garantire che i propri modelli siano accurati sia in fase di pre-addestramento che in fase di post-addestramento, nella misura in cui la tecnologia lo permette, al fine di renderli più utili per gli utenti.

In ultimo, con riferimento ai profili di competenza, il titolare ha affermato che a partire dal 15 febbraio 2024, OpenAI Ireland Ltd. agisce come parte contraente e fornitore di servizi per utenti e clienti localizzati nello SEE e in Svizzera, e come titolare o responsabile del trattamento. Pertanto, nel caso in cui fosse necessario affrontare i problemi di accuratezza dei modelli di OpenAI, coerentemente con le linee guida dell'EDPB, OpenAI avvierebbe ulteriori discussioni con l'IDPC, in qualità di autorità di controllo capofila.

3.1.6 CON RIFERIMENTO AL RISPETTO DELLE PRESCRIZIONI DEL GARANTE IMPARTITE CON IL PROVVEDIMENTO N. 114/2023.

Come sopra meglio illustrato, con il provvedimento n. 114 dell'11 aprile 2023, il Garante ha prescritto alcune misure correttive quali condizioni per la sospensione del provvedimento di limitazione provvisoria del 30 marzo 2023. L'ultima condizione, (punto 9 del provvedimento 114/2023), da adottare entro il 15 maggio 2023, riguarda la predisposizione di una campagna di informazione di natura non promozionale, su tutti i principali mezzi di comunicazione di massa italiani (radio, televisione, giornali e Internet) i cui contenuti andranno concordati con il Garante allo scopo di informare le persone dell'avvenuta probabile raccolta dei loro dati personali ai fini dell'addestramento degli algoritmi, dell'avvenuta pubblicazione sul sito Internet della società di un'apposita informativa di dettaglio e della messa a disposizione, sempre sul sito internet della società di uno strumento attraverso il quale tutti gli interessati potranno chiedere e ottenere la cancellazione dei propri dati personali".

Con nota del 15 maggio 2023 la Società ha reso noto – in asserito adempimento di tale prescrizione – di aver provveduto a: a) far pubblicare un'intervista della propria CTO sul quotidiano "La Repubblica"; b) acquistare una pagina intera su due quotidiani nazionali ove ha fatto pubblicare materiale educativo relativo al servizio ChatGPT; c) predisporre una pagina diretta agli utenti nel sito web di OpenAI; d) realizzare un video didattico da adottare in collaborazione con il Garante da rendere eventualmente pubblico anche tramite i canali istituzionali dell'Autorità.

L'Autorità, con nota del 18 maggio successivo, ha espresso "forte contrarietà" rispetto alla circostanza che la richiesta campagna mediatica fosse stata realizzata senza alcuna preventiva informazione e previo accordo con il Garante, riservandosi ogni valutazione rispetto al puntuale adempimento delle prescrizioni impartite con il richiamato provvedimento.

Alla luce del mancato corretto adempimento della misura correttiva prescritta l'Autorità ha contestato la violazione dell'art. 83, par. 5, lett. d), del Regolamento.

Nella memoria difensiva il titolare ha, tra l'altro, rappresentato che:

- OpenAI è desiderosa di collaborare con il Garante per quanto riguarda la campagna mediatica e apprezzerebbe molto l'opportunità di poter continuare il dialogo costruttivo su una nuova campagna di sensibilizzazione che soddisfi le aspettative del Garante;

- OpenAI ha pubblicato un gran numero di materiali informativi a partire dal 2016 e gestisce un indice di ricerca che comprende 170 post, documenti di ricerca, articoli e avvisi per informare le persone su vari argomenti relativi all'IA e allo sviluppo dei modelli;

- al momento dell'adozione del provvedimento n.114/2023, OpenAI non disponeva di un reparto marketing o di dipendenti specializzati nello sviluppo di campagne di marketing e non pubblicizzava o commercializzava i suoi prodotti o servizi, quindi non aveva esperienza nella organizzazione di campagne informative come quella prescritta dal Garante;

- in risposta al provvedimento n. 114/23, OpenAI ha lavorato rapidamente e in buona fede per organizzare e gestire la campagna informativa richiesta; in particolare ha informato il Garante dei contenuti della campagna informativa intrapresa e della disponibilità ad elaborarla ulteriormente;

- in sintesi, ha condotto in buona fede una campagna con modalità che riteneva adatte, appropriate e proporzionate agli elementi di contenuto richiesti dal Garante.

4. AUDIZIONE AI SENSI DELL'ART. 166, COMMA 6, D. LGS. N. 196/2003 E S.M.I., E DELL'ART. 13, COMMI 1 E 4, DEL REGOLAMENTO DEL GARANTE N. 1/2019

In data 11 aprile 2024, presso la sede del Garante per la Protezione dei dati personali, in piazza Venezia n. 11, si è tenuta l'audizione del titolare del trattamento nelle modalità di cui all'art. 13, comma 4, del Regolamento del Garante n. 1/2019 ed [OMISSIS].

In tale contesto la Società ha evidenziato come, sin dalla data del primo provvedimento, sia iniziata una fase di collaborazione ed interlocuzione con i vertici del Garante che ha condotto ad una lettera di impegni cui è seguito il secondo provvedimento dell'11 aprile 2023.

La Società ha altresì rimarcato l'impegno profuso in materia di protezione dei dati personali, sottolineando come le indicazioni del Garante abbiano consentito ad OpenAI di migliorare il funzionamento dei suoi modelli dal punto di vista della data protection e confermato la disponibilità di OpenAI di continuare a collaborare con l'Autorità.

Per quanto concerne la struttura societaria, la Società ha illustrato il suo percorso evolutivo,

rappresentando che OpenAI è nata nel 2015 come organizzazione di ricerca senza scopo di lucro per garantire lo sviluppo dell'intelligenza artificiale a beneficio di tutta l'umanità. Ai fini della realizzazione della missione della Società, è stata costituita in seno all'organizzazione una società di natura no profit nel 2023, che controlla l'intero gruppo e che determina l'indirizzo prevalente, una capped profit company, all'interno della quale gli investitori accedono ad un profitto limitato: il profitto che eccede il limite fissato viene reinvestito nell'organizzazione no profit.

Per quanto concerne i profili di trasparenza, la Società ha rappresentato di aver pubblicato una vasta gamma di materiali, inclusi schede informative sui sistemi, relazioni tecniche, articoli sul blog e altri lavori di ricerca e che, al momento del rilascio di ChatGPT nel novembre 2022, non aveva previsto il numero elevato di utenti che si è verificato, né di avere l'Italia come Paese target e che, ai tempi, l'interfaccia del servizio era disponibile solamente in lingua inglese. In merito alla trasparenza, la Società ha altresì specificato che la sua missione è sempre stata quella di condividere informazioni (ne sono un esempio le pubblicazioni sopra menzionate) ed ha precisato che l'informativa privacy esisteva già al momento del lancio di ChatGPT, pur dando atto che la stessa sia stata migliorata nel marzo 2023, prima dell'intervento del Garante, nonché successivamente, sulla base degli impegni concordati con il Garante. Inoltre la Società ha rappresentato di utilizzare un approccio multi-layer fornendo informazioni sia in maniera user-friendly, che più tecnica al fine di coprire un ampio spettro di utenti, da quelli dotati di maggiore conoscenza tecnica a quelli con meno esperienza.

Con riferimento alle risorse interne, la Società ha rappresentato che agli inizi del 2023 la Società contava meno di 400 dipendenti e di aver dovuto far fronte, in un arco di tempo brevissimo, a richieste di informazioni provenienti da molte Autorità di protezione dati.

La Società ha inoltre rappresentato che ChatGPT è stato rilasciato in buona fede e sin dall'inizio OpenAI ha adottato un modello di privacy compliance basato principalmente su tre elementi: la trasparenza, la privacy by design/il principio di minimizzazione e meccanismi di opt-out per gli utenti. Alla luce del grande e rapido successo riscosso da ChatGPT, la Società ha intrapreso numerose e continue attività di compliance tese al miglioramento del modello organizzativo (a titolo esemplificativo, la stesura della DPIA e della LIA). Parimenti, a seguito dell'intervento del Garante sono state implementate una serie di misure di compliance ulteriori, che oggi fanno parte dell'architettura della Società.

Rispetto alle misure di privacy by design, la Società ha precisato che ChatGPT non è un social network, non profila i suoi utenti, non tratta i dati personali a fini di marketing e che, sin dal rilascio di ChatGPT, ha reso disponibile agli interessati sistemi di opt-out per opporsi al trattamento di dati personali per finalità di post-training dei LLM sottesi al servizio.

La Società ha confermato che i dati personali trattati dai social network non vengono raccolti e trattati per finalità di addestramento dei modelli GPT nella misura in cui non sono pubblicamente disponibili (ad es. perché accessibili solamente previa registrazione di un account) e tali dati, in quanto tali, non sono ricompresi neppure nel dataset di Common Crawl a cui OpenAI ha attinto. Tali misure di minimizzazione erano già state implementate nella fase di addestramento del modello GPT3. Di contro, altre misure di minimizzazione indicate nella memoria difensiva afferiscono alla fase di post-training e sono oggetto di continuo aggiornamento e miglioramento.

L'Autorità ha chiesto ed ottenuto risposte anche in merito all'addestramento del modello GPT 3, che è stato effettuato negli USA in una fase precedente a quella del lancio di ChatGPT ed in cui, a parere della Società, non si applicava il Regolamento e che prima del rilascio di GPT 4, durante la fase di fine-tuning, è stata istituita una sandbox a cui hanno preso parte dei red teamers, ovverosia degli esperti, per migliorare gli aspetti relativi alla sicurezza del sistema tramite dei test finalizzati ad individuare risposte potenzialmente dannose.

L’Autorità ha chiesto e ottenuto riscontro anche in merito alla fase di post-training del modello, in relazione alla quale la Società ha precisato di aver implementato ulteriori misure di minimizzazione dei dati personali e di aver addestrato il modello a non restituire alcuna informazione di natura privata o sensibile a seguito di prompt degli utenti, nonostante le informazioni raccolte siano pubblicamente disponibili sul web. La Società ha quindi indicato nuovamente alcune delle misure di minimizzazione dei dati personali indicate nelle memorie difensive. Tra le tecniche di post-addestramento la Società ha, in particolare, segnalato di aver predisposto un intervento da parte di revisori umani - Reinforcement learning with human feedback (RLHF) al fine di ridurre la probabilità di risposte dannose o inesatte da parte dei LLM.

Tra le misure di minimizzazione è stato indicato anche il processo di deduplicazione dei dati in base al quale il modello apprende solo da informazioni che compaiono più volte sul web al fine di ridurre la possibilità di trattamento di dati personali riferiti a persone comuni.

La Società ha ribadito che il modello di ChatGPT non si basa su un database, ma consiste in stringhe di numeri e codici che interpretano e seguono i comandi e che tendono ad associare ad una parola quella successiva più probabile.

Quanto all’esercizio del diritto all’oblio, la Società ha specificato che esiste un apposito modulo in cui è possibile indicare le ragioni della richiesta, fermo restando il bilanciamento con eventuali altri diritti contrapposti. In particolare, è stato rappresentato che il diritto all’oblio viene garantito in due modi: 1) il dato di cui è stata chiesta la cancellazione non viene restituito nelle conversazioni del servizio; 2) il dato viene filtrato e non viene più utilizzato nell’addestramento dei modelli futuri. È stato altresì rappresentato che nel momento in cui il dato viene tokenizzato, ai fini dell’addestramento, deve considerarsi smaterializzato, e che le richieste di cancellazione da parte degli interessati sono attentamente monitorate da un team dedicato che rivede e tratta costantemente tali richieste.

Con riferimento al tema dell’esattezza dei dati personali è stato rappresentato che trattasi di una questione che riguarda in generale tutti i sistemi di intelligenza artificiale generativa e, su espressa richiesta dell’Autorità, è stato chiarito che, in presenza di notizie di stampa di cui sia stata chiesta la rettifica, la gestione dell’accuracy avviene informando gli utenti che, per ciascuna versione del modello, i dati sono raccolti fino ad una certa data. La Società sta inoltre valutando progetti di collaborazione con agenzie ed editori al fine di far sì che notizie attuali siano integrate nei servizi ChatGPT. Ad ogni modo è stato ricordato che l’utente viene avvertito dell’inesattezza del modello mediante un disclaimer nell’interfaccia utente. Inoltre quando l’inesattezza attiene a dati personali, la società dà seguito alle richieste di rettifica in due modi: ove possibile, tramite un’attività di post-addestramento specificamente finalizzata a correggere l’inesattezza; in alternativa, adottando un filtro nei dati di output di modo che il modello non restituisca tali informazioni.

Con riferimento all’esattezza, è stato altresì rappresentato che uno studio del novembre 2023 ha rilevato che ChatGPT vanta tra i più bassi tassi di allucinazione tra i principali servizi di intelligenza artificiale generativa.

L’Autorità ha chiesto e ottenuto precisazioni in merito alle versioni “turbo” di ChatGPT 3.5 e 4, che si differenziano dai rispettivi modelli “base” in quanto più aggiornate. È stato altresì chiarito che GPT 4 è stato implementato nel servizio ChatGPT Plus nel marzo 2023 e che ChatGPT Plus (un servizio a pagamento) è stato lanciato il 1° febbraio 2023 per ovviare ai problemi tecnici (crash) che si erano verificati su ChatGPT a causa dell’utilizzo massivo del servizio.

Con riferimento ai dati di pre-addestramento, come già chiarito nella memora difensiva, la Società ha precisato che la raccolta dei dati dal web è avvenuta quando il servizio non era ancora disponibile nell’Unione europea e che l’informativa in vigore nel novembre 2022, al lancio del servizio, già indicava il legittimo interesse come base giuridica per il miglioramento del servizio ed

era prevista la possibilità per gli utenti di esercitare l'opt-out. Di contro, la Società ha confermato che, inizialmente, non era previsto un sistema che consentisse l'esercizio dell'opt-out ai non utenti, sistema che è stato adottato dopo l'intervento del Garante.

Con riferimento alle violazioni contestate, la Società ha rappresentato di non aver ottenuto alcun vantaggio economico dalle stesse, che esse hanno avuto una durata limitata (novembre 2022 - marzo 2023) e che nel corso di tale periodo la Società non ha ricevuto alcun reclamo o istanza di esercizio dei diritti, per cui si presume che gli utenti non abbiano subito alcun danno.

Con riferimento ai dati di fatturato la Società ha riferito che lo stesso era di ridotte entità e che i ricavi sono iniziati a seguito del lancio della versione Plus. In particolare, nel periodo da gennaio a marzo 2023, i ricavi in Italia sono stati [OMISSIS]. Non ci sono stati utili distribuiti nel 2022 e nel 2023 e, [OMISSIS]. I ricavi sono stati reinvestiti nella fornitura e nel miglioramento del servizio, anche sotto il profilo privacy.

Pertanto è stato chiesto all'Autorità di tener conto di tali profili qualora decida di adottare un provvedimento sanzionatorio.

Da ultimo, con riferimento alla campagna informativa, la Società ha espresso rammarico ritenendo che sia occorso un misunderstanding con l'Autorità ed ha chiarito che, pur non disponendo di un ufficio comunicazione in Europa, sia all'epoca dei fatti che oggi, OpenAI si è attivata per elaborare una campagna informativa in un lasso di tempo limitato (11 aprile – 15 maggio 2023).

La Società ha ricordato che, ad ogni modo, la campagna è stata avviata, sebbene in assenza di un testo concordato con l'Autorità, in quanto i punti centrali della stessa risultavano ben delineati nel provvedimento n. 114/2023. È stata manifestata anche la disponibilità della Società di realizzare una nuova campagna informativa di concerto con il Garante attesa la manifestata volontà di cooperare con l'Autorità, nonostante OpenAI sia stabilita in Irlanda dal 15 febbraio 2024.

5. SUSSISTENZA DELLA GIURISDIZIONE EURO-UNITARIA E COMPETENZA DEL GARANTE

L'art. 3 del Regolamento disciplina l'Ambito di applicazione territoriale della normativa stabilendo criteri diversi a seconda che il titolare del trattamento sia o meno stabilito nel territorio dell'Unione europea.

Ai sensi dell'art. 3, par. 1, (c.d. criterio dello stabilimento), il Regolamento si applica indipendentemente dal fatto che il trattamento sia effettuato nell'Unione e la competenza viene individuata in ossequio al meccanismo del cd. sportello unico (one-stop-shop), ai sensi dell'art. 56 del Regolamento stesso.

Ai sensi dell'art. 3, par. 2 (c.d. criterio del targeting), il Regolamento si applica al trattamento dei dati personali di interessati che si trovano nell'Unione qualora le attività di trattamento riguardino: i) l'offerta di beni o la prestazione di servizi agli interessati nell'Unione (art. 3, par. 2, lett. a), del Regolamento); ii) il monitoraggio del comportamento di interessati che si trovano nell'Unione nella misura in cui tale comportamento abbia luogo nell'Unione stessa (art. 3, par. 2, lett. b), del Regolamento).

Nel caso di specie, all'epoca dei fatti contestati, OpenAI era stabilita in California, Stati Uniti d'America e non aveva individuato uno stabilimento nel territorio dell'Unione europea per cui, al fine di condurre una valutazione in ordine all'applicabilità della normativa europea in materia di protezione dei dati personali ai trattamenti posti in essere dalla Società, occorre verificare la sussistenza di uno dei due criteri di cui all'art. 3, par. 2, del Regolamento, sopra richiamati.

Atteso che il servizio ChatGPT è stato reso disponibile al pubblico il 30 novembre 2022, occorre accertare se detto servizio possa considerarsi offerto ad interessati che si trovano nell'Unione

europea per l'applicabilità del criterio del targeting di cui alla lettera a) del citato art. 3 del Regolamento e da quale data.

Al riguardo, si richiamano le Guidelines 3/2018 on territorial scope, adottate dal Comitato per la protezione dei dati personali il 12 novembre 2019, le quali prevedono che il "titolare del trattamento… dimostr[i] la sua intenzione di offrire beni o servizi a un interessato che si trova nell'Unione" (cfr. par. 2.a delle Linee guida citate) e la giurisprudenza della Corte di Giustizia dell'Unione europea (sentenza Pammer/Reederei Karl Schlüter GmbH & Co e Hotel Alpenhof/Heller (cause riunite C-585/08 e C-144/09), la quale ha indicato alcuni fattori in presenza dei quali possa ritenersi che un'attività commerciale svolta da un soggetto sia diretta nei confronti di uno Stato membro, citando, tra gli altri, la circostanza che l'Unione europea sia menzionata in riferimento al bene o servizio offerto, la natura internazionale dell'attività oppure l'avvio di campagne pubblicitarie e di marketing rivolte al pubblico di un paese dell'UE.

Nel caso di specie, si osserva innanzitutto che, nel corso del procedimento, la Società stessa ha dichiarato che Prima del 30 novembre 2022 OpenAI non aveva sede nell'UE, non offriva beni o servizi agli interessati localizzati nell'UE e non monitorava il loro comportamento (memoria difensiva, versione tradotta, pag. 18), riconoscendo come dies a quo la data del 30 novembre 2022.

La Società ha tentato di posticipare ad un momento indeterminato, successivo al 30 novembre 2022, l'applicabilità del Regolamento sostenendo che Una volta che gli interessati localizzati nell'UE lo hanno scoperto e utilizzato come servizio, il GDPR, compreso l'obbligo di identificare una base giuridica e completare una LIA, è diventato applicabile (memoria difensiva, versione tradotta, pag. 19), tuttavia dall'interpretazione della norma risulta pacifico che l'offerta del servizio deve essere ricondotta ad un'azione del titolare (intenzione di offrire un servizio) e non alla "scoperta" del servizio stesso da parte degli interessati. In tal senso, si rileva che ChatGPT è stato reso disponibile gratuitamente online come anteprima di ricerca gratuita costituita da una semplice interfaccia progettata per facilitare l'interazione delle persone con la ricerca di OpenAI (cfr. allegato al verbale di audizione, pag. 3). L'intenzione di OpenAI di offrire un servizio a livello globale si evince dunque dalle caratteristiche del servizio stesso (online, gratuito e user-frendly) nonché dal fatto che, a distanza di poco più di un mese, L'11 gennaio 2023, OpenAI e i suoi consulenti legali esterni hanno iniziato a lavorare sulla documentazione della LIA come parte della valutazione d'impatto sulla protezione dei dati (cfr. memoria difensiva, versione tradotta, pag. 19).

Considerato che la data del 30 novembre corrisponde alla data del rilascio al pubblico del servizio ChatGPT e che tale servizio era pacificamente disponibile anche per interessati che si trovavano nell'Unione europea, può ritenersi integrata l'ipotesi di cui all'art. 3, par. 2, lett. a), del Regolamento e dunque la sussistenza della giurisdizione euro-unitaria a decorrere da tale data.

Quanto dalla competenza del Garante, si osserva quanto segue.

Il trattamento di dati personali posto in essere da OpenAI è qualificabile come trattamento transfrontaliero di dati personali ai sensi dell'art. 4, par. 1, n. 23 del Regolamento in quanto idoneo ad incidere su interessati in più di uno Stato membro.

Per questa tipologia di trattamenti, laddove il titolare abbia individuato uno stabilimento, unico o principale, nell'Unione europea, trova applicazione il meccanismo di cooperazione descritto negli artt. 60 ss. del Regolamento in base al quale la competenza ad esercitare i compiti ed i poteri di cui agli artt. 57 e 58 del Regolamento è radicata, ai sensi dell'art. 56, par. 1, del Regolamento in capo alla autorità di controllo capofila, ovverosia l'autorità di controllo dello Stato membro in cui si trova lo stabilimento unico o principale.

Qualora, al contrario, non esista nel territorio europeo uno stabilimento del titolare del trattamento,

quest'ultimo dovrà interfacciarsi con le autorità di controllo di ciascuno Stato membro in cui opera per il tramite del rappresentante designato (cfr. par. 3.3 delle "Guidelines on the Lead Supervisory Authority" adottate dal Gruppo di Lavoro Articolo 29 il 13 dicembre 2016, revisionate il 5 aprile 2017 e fatte proprie dal Comitato per la protezione dei dati personali in data 25 maggio 2018).

Infatti, laddove un titolare non disponga di uno stabilimento nell'Unione europea (rectius nell'area SEE), la norma speciale di cui all'art. 56 non trova applicazione a favore della regola generale di cui all'art. 55, par. 1, del Regolamento secondo cui ogni Autorità di controllo è compente ad eseguire i compiti assegnati ed a esercitare i poteri ad essa conferiti a norma del (…) regolamento nel territorio del rispettivo Stato membro.

Nel caso in esame, come detto, OpenAI è una società con sede negli Stati Uniti d'America che, fino al 15 febbraio 2024, non aveva uno stabilimento nel territorio dell'Unione europea e, pertanto, l'autorità di protezione dei dati personali italiana è competente a valutare, con riguardo al proprio territorio, la conformità al Regolamento del trattamento di dati personali posto in essere da OpenAI fino a tale data e ad esercitare i poteri ad essa riconosciuti dall'art. 58.

A tal riguardo, si rileva che l'EDPB, nel Parere n. 8/2019 sulla competenza di un'autorità di controllo in caso di mutamento delle circostanze relative allo stabilimento principale o unico, del 9 luglio 2019, ha precisato che il cambiamento del ruolo di autorità di controllo capofila non significa che l'autorità di controllo iniziale non fosse competente ad agire nel momento in cui lo ha fatto e, pertanto, tale circostanza non priva retroattivamente di base giuridica le operazioni già svolte dall'autorità capofila iniziale. L'autorità di controllo precedentemente competente aveva piena giurisdizione quando lo stabilimento principale o unico era situato sul suo territorio. Di conseguenza, gli atti compiuti conservano il loro valore e gli elementi di prova e le informazioni raccolti dalla precedente autorità di controllo capofila possono essere utilizzati da quella divenuta successivamente competente. Il Comitato ha ritenuto applicabile lo stesso principio, nel medesimo parere, all'ipotesi in cui un titolare costituisca uno stabilimento principale o unico nell'area SEE, specificando che la competenza dell'autorità non capofila viene meno solo nel momento in cui tale mutamento di circostanze diventa efficace ed è comprovato e che qualsiasi procedimento amministrativo pendente (necessariamente un procedimento non soggetto a cooperazione data l'assenza iniziale di uno stabilimento principale nell'area SEE) deve essere trasferito all'autorità di controllo dello Stato nel quale si trova lo stabilimento principale.

Nel caso di specie, il procedimento amministrativo pendente avanti il Garante non può essere tout court traferito all'autorità di controllo capofila irlandese, competente ai sensi dell'art. 56, par. 1, del Regolamento, alla data di adozione del presente provvedimento, in quanto alcune delle violazioni accertate sono state consumate in data anteriore al 15 febbraio 2024. I principi sopra richiamati desunti dal parere n. 8/2019 si applicano, infatti, esclusivamente a questioni di competenza che vertono su violazioni di natura continuativa o continuata. Secondo la Corte europea dei diritti dell'uomo, Grande Camera, con violazione "continuativa" si intende un'azione (o un'omissione) che dura per un certo lasso di tempo mentre con violazione "continuata" si intendono più azioni che contengono tutti gli elementi dello stesso (o di un analogo) atto illecito commessi nell'arco di un certo periodo di tempo (cfr. sentenza della Grande Camera, 18 aprile 2013, nella causa Rohlena/Repubblica ceca, n. 59552/08).

Nell'ipotesi di trattamenti transfrontalieri, il trasferimento di un procedimento amministrativo da un'autorità di controllo (che ha un'istruttoria aperta al momento della costituzione o del trasferimento di uno stabilimento del titolare) all'autorità capofila avviene, dunque, solo nel caso in cui la violazione sia ancora in corso (vuoi perché si tratta di una violazione continuativa, vuoi perché si tratta di una violazione continuata).

Una conferma di tale interpretazione è ravvisabile anche nelle linee guida sul calcolo della sanzione (cfr. Guidelines 04/2022 on the calculation of administrative fines under the GDPR,

Versione 2.1., 24 Maggio 2023), nella parte in cui si fa riferimento al principio di pluralità delle azioni (concorso materiale) per descrivere le fattispecie non sussumibili nell'art. 83, par. 3, del GDPR (violazione continuata). In tali ipotesi le sanzioni erogate per ogni singola violazione si sommano materialmente in quanto il titolare ha agito mediante distinte plurime azioni.

Nell'ipotesi di una pluralità di azioni poste in essere da un titolare del trattamento in violazione del Regolamento, la competenza a decidere può, dunque, essere radicata anche in capo a più di una autorità di controllo, nell'ipotesi in cui, medio tempore, il titolare abbia costituito o trasferito uno stabilimento nell'Unione europea, a condizione che si tratti di violazioni consumate e di natura non continuata, alla data dell'effettività del nuovo stabilimento. In tal caso l'esercizio del potere sanzionatorio (segnatamente l'erogazione di plurime sanzioni per plurime violazioni integrate mediante distinte, plurime azioni) da parte dell'autorità di controllo competente prima della costituzione o del trasferimento dello stabilimento è legittimo, fermo restando l'obbligo di rispettare il principio generale di proporzionalità di cui all'art. 83, par. 1, del Regolamento.

Nel caso di specie, dunque, deve ritenersi radicata la competenza del Garante per ogni violazione del Regolamento da parte di OpenAI consumata prima del 15 febbraio 2024 e di natura non continuata.

6. LE VIOLAZIONI ACCERTATE

6.1 ART. 33 DEL REGOLAMENTO

L'Ufficio ha contestato ad OpenAI la violazione dell'art. 33 del Regolamento per aver omesso di notificare all'Autorità l'evento di violazione dei dati occorso il 20 marzo 2023.

L'art. 33, del Regolamento, prevede che in caso di violazione dei dati personali, il titolare del trattamento notifica la violazione all'autorità di controllo competente a norma dell'articolo 55 senza ingiustificato ritardo e, ove possibile, entro 72 ore dal momento in cui ne è venuto a conoscenza, a meno che sia improbabile che la violazione dei dati personali presenti un rischio per i diritti e le libertà delle persone fisiche. Qualora la notifica all'autorità di controllo non sia effettuata entro 72 ore, è corredata dei motivi del ritardo.

Dall'istruttoria è emerso che il 20 marzo 2023, alle ore 7:50 (fuso orario del Pacifico), OpenAI ha ricevuto da un utente la segnalazione di una anomalia nel funzionamento del servizio ChatGPT. L'incidente di sicurezza, non causato da attacchi esterni, ha riguardato la confidenzialità delle chat di alcuni utenti che sono state visibili, con indicazione dei relativi dati di registrazione, ad altri utenti in una finestra temporale compresa tra le ore 01:00 e le ore 10:00 (fuso orario del Pacifico).
Dagli atti è emerso che la causa dell'incidente è da attribuire ad un problema di configurazione dei sistemi di OpenAI che sono progettati per distribuire il traffico in entrata e in uscita (ossia, l'insieme dei prompt degli utenti e le relative risposte generate dal programma) in accordo a un criterio di bilanciamento di carico (workload management). Il meccanismo di distribuzione e accodamento dei prompt e delle relative risposte (in modo che ogni utente riceva la corretta sequenza di risposte dal servizio) avviene sotto la supervisione [OMISSIS]. A causa di errori presenti in alcune righe di codice, per talune coppie prompt-risposta è stata effettuata una duplicazione del traffico e una ricongiunzione del flusso su code distinte riferibili a utenti diversi, con l'effetto della perdita di confidenzialità sopra richiamato. La Società è intervenuta per correggere gli errori del codice e introdurre ulteriori controlli (redundancy checks) in modo che simili eventi non potessero ripetersi in futuro. Dall'analisi dei file di log la Società è risalita alle identità degli utenti potenzialmente coinvolti.

Il 23 marzo 2023, nel termine di 72 ore previsto dall'art. 33 del Regolamento, OpenAI ha notificato il data breach all'Autorità di controllo irlandese, ritenendo che la stessa avrebbe condiviso le informazioni con gli altri membri del Comitato tra cui il Garante. OpenAI ha riferito di aver incluso

nella notifica il numero di soggetti potenzialmente coinvolti nell'Unione europea e la loro ripartizione a livello nazionale (440 in Italia).

Il 24 marzo 2023, sebbene il rischio per i diritti e le libertà delle persone coinvolte nel data breach fosse stato valutato basso, la Società ha volontariamente intrapreso l'ulteriore iniziativa di pubblicare un post sul proprio sito web e di informare tutti gli utenti potenzialmente coinvolti (cfr. memoria difensiva, versione tradotta, pag. 11).

Il 27 aprile 2023, a fronte di una richiesta di chiarimenti del 13 aprile 2023, OpenAI ha fornito dall'Autorità irlandese ulteriori informazioni sull'incidente.

Le argomentazioni addotte dalla Società per giustificare l'avvenuta notifica all'Autorità irlandese, che fanno leva sulla costituzione della società OpenAI Ireland Ldt. avvenuta a distanza di pochi giorni dal data breach, non possono trovare accoglimento.

Come illustrato nel paragrafo precedente, infatti, il meccanismo dello sportello unico non era applicabile al tempo in cui è si è verificato l'incidente di sicurezza in quanto la Società non era stabilita nell'Unione europea. Come chiarito dal Comitato con Parere n. 8/2019 sulla competenza di un'autorità di controllo in caso di mutamento delle circostanze relative allo stabilimento principale o unico, del 9 luglio 2019, il concetto di stabilimento non può essere ridotto ad un mero atto estemporaneo o burocratico, trattandosi, al contrario, di un'operazione concreta, attuata secondo propositi duraturi. Peraltro, anche volendo avallare tale tesi, si osserva come, per stessa ammissione della Società, alla data del data breach la società irlandese non fosse ancora stata costituita.

Nel caso di specie, dunque, la Società avrebbe dovuto notificare la violazione dei dati ex art. 33 del Regolamento a tutte le autorità europee di protezione dati i cui interessati erano stati coinvolti nel data breach. Atteso che dagli atti istruttori è emerso che l'evento ha interessato 440 utenti italiani, OpenAI avrebbe dovuto notificare la violazione dei dati direttamente al Garante.

Nel merito, si rileva che l'incidente ha riguardato l'1,2% degli utenti del servizio a pagamento ChatGPT Plus, attivi nel corso della finestra temporale indicata; l'incidente non ha dunque riguardato la generalità degli utenti di ChatGPT, ma una esigua percentuale dei soli utenti a pagamento attivi in una specifica finestra temporale di nove ore e che la società successivamente ha approntato misure di sicurezza volte a impedire il reiterarsi di simili incidenti.

Per quanto sopra rappresentato, si ritiene che OpenAI abbia violato l'art. 33, par. 1. del Regolamento e che detta violazione debba ritenersi consumata il 23 marzo 2023 (72 ore dopo il data breach) e sia di natura non continuata.

6.2 ARTT. 5, PAR. 2 E 6 DEL REGOLAMENTO

L'Ufficio ha contestato ad OpenAI la violazione degli artt. 5, par. 2 e 6 del Regolamento per non essere stata in grado di dimostrare di aver chiaramente individuato sino alla data del 30 marzo 2023 e, in ogni caso, prima dell'inizio dell'attività di trattamento, una base giuridica per il trattamento dei dati personali per finalità di addestramento del modello GPT sotteso al funzionamento del servizio ChatGPT, reso disponibile al pubblico a far data dal 30 novembre 2022.

L'art. 5, par. 1, del Regolamento prescrive che i dati personali sono: a) trattati in modo lecito, corretto e trasparente nei confronti dell'interessato («liceità, correttezza e trasparenza»); b) raccolti per finalità determinate, esplicite e legittime, e successivamente trattati in modo che non sia incompatibile con tali finalità; un ulteriore trattamento dei dati personali a fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici non è, conformemente all'articolo 89, paragrafo 1, considerato incompatibile con le finalità iniziali («limitazione della

finalità»); c) adeguati, pertinenti e limitati a quanto necessario rispetto alle finalità per le quali sono trattati («minimizzazione dei dati»); d) esatti e, se necessario, aggiornati; devono essere adottate tutte le misure ragionevoli per cancellare o rettificare tempestivamente i dati inesatti rispetto alle finalità per le quali sono trattati («esattezza»); e) conservati in una forma che consenta l'identificazione degli interessati per un arco di tempo non superiore al conseguimento delle finalità per le quali sono trattati; i dati personali possono essere conservati per periodi più lunghi a condizione che siano trattati esclusivamente a fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici, conformemente all'articolo 89, paragrafo 1, fatta salva l'attuazione di misure tecniche e organizzative adeguate richieste dal presente regolamento a tutela dei diritti e delle libertà dell'interessato («limitazione della conservazione»); f) trattati in maniera da garantire un'adeguata sicurezza dei dati personali, compresa la protezione, mediante misure tecniche e organizzative adeguate, da trattamenti non autorizzati o illeciti e dalla perdita, dalla distruzione o dal danno accidentali («integrità e riservatezza»). Il paragrafo 2 della stessa norma prevede che Il titolare del trattamento è competente per il rispetto del paragrafo 1 e in grado di comprovarlo («responsabilizzazione»).

La responsabilizzazione del titolare, ai sensi dell'art. 5, par. 2, del Regolamento, ha una duplice valenza: da un lato, mira a spingere il titolare ad adottare misure utili a prevenire e limitare gli impatti che le attività di trattamento possano avere sulla sfera giuridica degli interessati, in termini di gravità e di probabilità, dall'altro, persegue il fine di documentare quanto fatto in chiave probatoria, a fronte di violazioni effettive o comunque di controlli dell'Autorità.

L'art. 6 del Regolamento prescrive le condizioni di liceità del trattamento elencando le sei possibili basi giuridiche (consenso, contratto, obbligo di legge, interesse vitale, interesse pubblico, interesse legittimo) su cui il titolare deve fare affidamento per poter trattare lecitamente i dati personali necessari per lo svolgimento della propria attività. La base giuridica, come chiarito dall'EDPB, deve essere individuata prima dell'attuazione del trattamento e deve essere specificata nelle informazioni fornite agli interessati conformemente agli articoli 13 e 14. (cfr. le Linea guida 2/2019 sul trattamento di dati personali ai sensi dell'articolo 6, paragrafo 1, lettera b), del regolamento generale sulla protezione dei dati nel contesto della fornitura di servizi online agli interessati").

In sede difensiva la Società ha rappresentato che, alla data del lancio del servizio ChatGPT, in data 30 novembre 2022, OpenAI era in linea con i principi di protezione dei dati, ma non era soggetta al GDPR (cfr. memoria difensiva, versione tradotta, pag. 19).

Per contro, come già evidenziato nel paragrafo 5 del presente atto, l'applicabilità del Regolamento è riconducibile alla data del rilascio del servizio e non ad un indeterminato momento successivo. Da quanto risulta in atti, tuttavia, a tale data (30 novembre 2022) - sebbene il servizio offerto fosse idoneo a produrre un impatto significativo sui diritti e sulle libertà degli interessati, in ragione sia dell'innovatività e della complessità della tecnologia utilizzata (LLM) che del numero degli interessati potenzialmente coinvolti - OpenAI non aveva formalizzato l'identificazione della base giuridica del trattamento dei dati personali per fini di addestramento del modello linguistico GPT sotteso al funzionamento del servizio ChatGPT. In particolare, sebbene la Società abbia sostenuto di aver basato le operazioni di trattamento relative alla fornitura del servizio agli utenti sulla base giuridica dell'esecuzione del contratto e le operazioni di trattamento relative all'addestramento degli algoritmi sulla base giuridica del legittimo interesse, dalla documentazione prodotta non risulta che tale indicazione fosse stata formalizzata alla data del 30 novembre 2022. Segnatamente, la Società ha depositato, rispettivamente allegate alle note del 19 maggio e 20 novembre 2023, una copia della DPIA e della LIA: la prima risulta redatta in bozza il 24 febbraio 2023 ed aggiornata il 19 maggio 2023, mentre la seconda è stata prodotta, solo a fronte di espressa richiesta dell'Autorità, priva di data. Con nota del 23 aprile 2024, a scioglimento di una riserva (cfr. verbale audizione del 19 aprile24 pag.6) assunta in sede di audizione, OpenAI ha rappresentato che la prima bozza di LIA è stata completata in data 24 febbraio 2023, quindi ben

dopo l’inizio del trattamento (30 novembre 2022).

L’Autorità ritiene che tali documenti non consentano di dimostrare che la valutazione di adeguatezza e l’individuazione della base giuridica per il trattamento di dati personali per finalità di addestramento del modello linguistico GPT sotteso al funzionamento del servizio ChatGPT sia avvenuta prima del lancio del servizio stesso (30 novembre 2022) come richiesto dal principio di accountability di cui all’art. 5, par. 2 e dall’art. 6 del Regolamento.

Si rileva, inoltre, che la circostanza che il predetto servizio sia stato immesso sul mercato senza formalizzare alcuna base giuridica per il trattamento di dati personali e senza aver fornito idonee informazioni agli utenti (cfr. successivo paragrafo 6.3), abbia di fatto reso impossibile agli interessati l’esercizio dei loro diritti, in violazione del principio di accountabilty.

Ciò posto, si conferma quanto affermato in sede di contestazione in merito all’incapacità della documentazione in atti di far emergere elementi da cui far discendere la prova che la Società abbia individuato la base giuridica del trattamento dei dati per fini di addestramento del modello linguistico GPT in un momento anteriore all’inizio dello stesso. Tale carenza risulta altresì confermata dall’assenza di qualsivoglia indicazione in merito alla condizione di liceità relativa al trattamento di dati personali per finalità di addestramento del modello GPT nelle informazioni fornite agli interessati (sul punto, cfr. paragrafo successivo). A tal proposito non pare dirimente neppure l’argomento difensivo secondo cui sarebbe esaustivo il riferimento, nella privacy policy aggiornata al 14 marzo 2023, al legittimo interesse per lo “sviluppo dei servizi”. Tale indicazione, infatti, oltre che tardiva rispetto all’inizio del trattamento coincidente con il lancio del servizio sul mercato effettuato in novembre, non si ritiene sufficientemente chiara ed intelligibile, a prescindere da ogni valutazione di merito relativa all’idoneità del legittimo interesse quale base giuridica del trattamento e considerato che la stessa risulterebbe di competenza della Autorità di controllo irlandese, in quanto relativa ad una violazione di natura continuata

In particolare si rileva che OpenAI nell’identificare il legittimo interesse quale base giuridica del trattamento, specie nel caso dei non utenti, abbia omesso di esplicitare nell’informativa che il trattamento si basava sull'art. 6, par. 1, lett. f), del Regolamento e quali erano i legittimi interessi perseguiti, in tal modo compromettendo la possibilità da parte degli interessati di poter esercitare il diritto di opposizione ai sensi dell’art. 21 del Regolamento.

Per quanto sopra rappresentato, si ritiene che OpenAI non sia stata in grado di dimostrare di aver identificato una base giuridica prima di iniziare le attività di trattamento dei dati personali in violazione degli articoli 5, par.2, e 6 del Regolamento. Detta violazione, accertata alla data del 30 marzo 2023, deve ritenersi consumata il 30 novembre 2022 e sia di natura non continuata.

Per quanto concerne le valutazioni in merito alla legittimità della scelta compiuta successivamente all’inizio del trattamento e al lancio del relativo servizio in ordine al legittimo interesse come base giuridica per il trattamento dei dati personali di utenti e non utenti per l’addestramento degli algoritmi e il funzionamento del servizio, l’Autorità, conformemente alle citate regole sul riparto di giurisdizione a seguito dello stabilimento di OpenAI in Irlanda, trasmette all’Autorità di controllo irlandese, autorità capofila ai sensi dell’art. 56 del Regolamento, gli atti del presente procedimento per la prosecuzione delle attività di sua competenza.

6.3 ARTT. 5, PAR. 1, LETT. A), 12 E 13 DEL REGOLAMENTO

L’Ufficio ha contestato ad OpenAI la violazione degli artt. 5, par. 1, lett. a), 12 e 13 del Regolamento per omissioni e carenze nella privacy policy (versione del 14 marzo 2023), vigente al 30 marzo 2023.

L’art. 5, par. 1, lett. a), del Regolamento prescrive che i dati personali siano trattati in modo lecito,

corretto e trasparente nei confronti dell'interessato (principio di liceità, correttezza e trasparenza). L'art. 12 del Regolamento detta norme in materia di trasparenza e modalità di esercizio dei diritti, mentre l'art. 13 introduce indicazioni specifiche in ordine alle informazioni che il titolare è tenuto a fornire qualora i dati personali siano raccolti presso l'interessato.

In tema di trasparenza il Considerando 58 del Regolamento prevede che le informazioni destinate al pubblico o all'interessato siano concise, facilmente accessibili e di facile comprensione, che sia usato un linguaggio semplice e chiaro e, con riferimento alla protezione specifica di cui debbano essere destinatari i minori, prevede che quando il trattamento dati li riguarda, qualsiasi informazione e comunicazione dovrebbe utilizzare un linguaggio semplice e chiaro che un minore possa capire facilmente.

Sul tema della trasparenza rilevano altresì le indicazioni del Comitato, in particolare le linee guida 2/2019 sul trattamento dei dati personali ai sensi dell'articolo 6, par.1, lett. b) del regolamento generale sulla protezione dei dati nel contesto della fornitura di servizi online agli interessati, ove è previsto che la base giuridica del trattamento, oltre a dover essere individuata prima dell'attuazione del trattamento, deve essere specificata nelle informazioni fornite agli interessati conformemente agli articoli 13 e 14; rilevano, inoltre le linee guida n.1/2022 del Comitato in materia di accesso, che prescrivono, al paragrafo 142, che un titolare del trattamento che offre un servizio in un determinato paese dovrebbe anche rispondere in una lingua compresa dagli interessati di quel paese.

L'istruttoria dell'Ufficio ha avuto ad oggetto la privacy policy adottata e pubblicata da OpenAI alla data del 30 marzo 2023, ovverosia la versione della stessa aggiornata al 14 marzo 2023, segnatamente le informazioni rese agli interessati con riferimento sia ai dati degli utenti del servizio richiesti dalla Società per l'accesso e l'utilizzo della piattaforma, sia ai dati disponibili in rete, relativi ad utenti e non-utenti, trattati da OpenAI nell'ambito delle attività di addestramento dei modelli GPT 3.5 e GPT 4, sottesi rispettivamente al funzionamento dei servizi ChatGPT (servizio gratuito) e ChatGPT Plus (servizio a pagamento).

Dall'istruttoria è innanzitutto emerso che, alla data del 30 marzo 2023, la privacy policy della Società era disponibile solo in lingua inglese (anche per i minori) e non risultava di facile reperibilità nel sito; in particolare, il link presente nel flusso di registrazione era collocato in una posizione tale da non consentirne agli utenti la lettura prima di procedere all'inserimento dei dati per la creazione di un account.

Inoltre, le informazioni rese al tempo dal titolare si riferivano, esclusivamente, ai dati degli utenti trattati dalla Società per l'utilizzo del servizio, mentre, sia agli utenti che ai non utenti, non veniva fornita alcuna informazione in relazione al trattamento dei dati personali per fini di addestramento degli LLM.

Le argomentazioni difensive del titolare non hanno consentito di superare le valutazioni sostenute dall'Ufficio in sede di contestazione.

La Società, nella memoria difensiva, ha sostenuto di aver assicurato, alla data del 30 marzo 2023, trasparenza agli utenti del servizio ChatGPT sul trattamento delle loro conversazioni per la messa a punto dei modelli GPT e ai non-utenti sul trattamento delle informazioni disponibili al pubblico per l'addestramento dei modelli.

Quanto al primo profilo, OpenAI ha richiamato il testo della privacy policy, vigente all'epoca della contestazione, nella parte relativa agli User Content le cui finalità erano Fornire, amministrare, mantenere, migliorare e/o sviluppare i Servizi, un pop-up, presente sin dal lancio del servizio, che avvisava gli utenti che le loro conversazioni avrebbero potuto essere esaminate da addestratori di IA per "migliorare i … sistemi", nonché due articoli, aggiornati al 1° marzo 2023, pubblicati

nell'Help Center.

Si osserva come il linguaggio utilizzato nella privacy policy non sia affatto chiaro, atteso che la locuzione miglioramento/sviluppo del servizio non rende facilmente associabile la specifica, peculiare ed innovativa finalità di addestramento dei modelli di intelligenza artificiale generativa sottesi a ChatGPT (da intendersi nel senso di affinamento - fine-tuning), alla generica finalità di miglioramento del servizio, usualmente perseguita dai fornitori di servizi online. Parimenti poco chiara e non intuitiva è la riconducibilità della finalità di ricerca alla ricerca specifica nel campo dell'intelligenza artificiale (cfr. memoria difensiva, versione tradotta, pag. 25). Sotto altro profilo, si rileva come l'adozione di un pop-up e la pubblicazione di due articoli pubblicati nell'Help Center non siano suscettibili di integrare l'adempimento degli obblighi di cui agli artt. 12 e 13 del Regolamento, atteso che le informazioni agli utenti devono innanzitutto essere previste nella privacy policy e solo in misura aggiuntiva, sebbene apprezzabile, in altri documenti. Da ultimo, si osserva che le informazioni agli utenti riguardavano, comunque, solamente il trattamento dei dati personali connesso all'utilizzo del servizio e non quelli trattati nella fase di addestramento dei modelli.

Quanto alle argomentazioni difensive relative al presunto adempimento del principio di trasparenza nei confronti dei non-utenti, la Società ha affermato di aver adottato molteplici misure per informare gli utenti e i non utenti del trattamento per fini di addestramento del modello GPT, ed a conferma di ciò ha citato la pubblicazione, sin dal 2019, di numerosi documenti di ricerca, schede tecniche, articoli e post che illustravano in dettaglio l'uso da parte di OpenAI di dati Internet pubblici per l'addestramento dei suoi modelli linguistici (cfr. memoria difensiva, versione tradotta, pag. 26). Al riguardo si osserva come la Società non abbia chiarito come intendesse garantire la conoscibilità di tali informazioni da parte degli interessati. Le informazioni di cui agli artt. 12 e 13 del Regolamento costituiscono il principale strumento individuato dal legislatore europeo per rendere edotti gli interessati dei trattamenti che coinvolgono i loro dati personali, salvo eccezionali casi espressamente previsti dalla legge. In tal senso, è evidente che la documentazione citata, sebbene molto varia ed anche tecnicamente di rilievo, non possa sopperire all'obbligo di trasparenza ai sensi del Regolamento atteso che non è dato comprendere la ragione per cui gli interessati, utenti e – soprattutto - non utenti, avrebbero dovuto accedere a tali documenti informativi e prenderne conoscenza. Inoltre, anche ammesso e non concesso che tali documenti fossero ritenuti idonei all'adempimento dell'obbligo, non si comprende la ragione per cui gli interessati avrebbero potuto ragionevolmente attendersi che i loro dati pubblicamente disponibili avrebbero potuto essere trattati ai fini di addestramento dei modelli di OpenAI.

Tali argomentazioni difensive risultano altresì contraddittorie rispetto all'assunto difensivo secondo cui il successo del servizio è stato dirompente quanto inaspettato. Infatti, proprio in ragione dell'ambito circoscritto ed iper specializzato in cui operava inizialmente il titolare, si ritiene particolarmente improbabile che un utente medio accedesse e consultasse, sua sponte, i documenti di ricerca, le pubblicazioni, gli articoli e le comunicazioni asseritamente contenenti, tra le altre informazioni, anche di carattere tecnico, informazioni sul trattamento dei loro dati personali.

Con riferimento all'art. 5, par. 1, lett. a), del Regolamento, si richiama in questa sede il principio espresso dall'EDPB nella decisione vincolante n. 1/2021, secondo cui la trasparenza deve essere considerata un concetto generale che trova concreta attuazione in diverse disposizioni e obblighi specifici (ad esempio, gli artt. 12, 13, 14, 25 e 35 del Regolamento). È, dunque, necessario distinguere gli obblighi specifici derivanti dal principio di trasparenza dal principio stesso espresso nell'art. 5 del Regolamento, in quanto il primo non può essere circoscritto agli obblighi di cui agli artt. 12-14del Regolamento, sebbene questi ultimi siano una concretizzazione del precedente. Il principio di trasparenza, infatti, è un principio onnicomprensivo che rafforza altri principi (es. correttezza, accountability). La conferma di tale ricostruzione è data dal fatto che l'art. 83, par. 5, del Regolamento prevede la possibilità di sanzionare la violazione degli obblighi di trasparenza indipendentemente dalla violazione del principio stesso. Nel caso di specie, si ritiene ravvisabile

anche la violazione del principio di trasparenza di cui all'art. 5, par. 1, lett. a) del Regolamento alla luce della gravità della carenza informativa sia per gli utenti che per i non-utenti, i quali non erano al corrente del trattamento dei loro dati personali e non potevano ragionevolmente aspettarsi un trattamento dei loro dati. Infatti, come già rappresentato nel paragrafo precedente, mancava qualsivoglia informazione in relazione alla base giuridica del trattamento dei dati personali per finalità di addestramento dei modelli GPT, alla natura del trattamento (mancanza di informazioni circa gli elementi essenziali del trattamento come la base giuridica) e all'impatto dello stesso (alla luce della innovativa tipologia di trattamento connessa alla tecnologia in argomento).

Per quanto sopra rappresentato, si ritiene che OpenAI abbia violato gli artt. 5, par. 1, lett. a), 12 e 13 del Regolamento e che detta violazione, accertata alla data del 30 marzo 2023, debba ritenersi consumata il 30 novembre 2022 e sia di natura non continuata.

6.4 ARTT. 24 E 25, PAR. 1, DEL REGOLAMENTO

L'Ufficio ha contestato ad OpenAI la violazione degli artt. 8, 24 e 25, par. 1, del Regolamento per omessa predisposizione di idonei sistemi atti a verificare l'età dei soggetti alla data del 30 marzo 2023.

Ai sensi dell'art. 24, par.1, del Regolamento Tenuto conto della natura, dell'ambito di applicazione, del contesto e delle finalità del trattamento, nonché dei rischi aventi probabilità e gravità diverse per i diritti e le libertà delle persone fisiche, il titolare del trattamento mette in atto misure tecniche e organizzative adeguate a garantire, ed essere in grado di dimostrare, che il trattamento è effettuato conformemente al presente regolamento. Dette misure sono riesaminate e aggiornate qualora necessario.

Ai sensi dell'art. 25, par. 1, del Regolamento, il titolare deve adottare tali misure tenendo conto dello stato dell'arte e dei costi di attuazione, nonché della natura, dell'ambito di applicazione, del contesto e delle finalità del trattamento, come anche dei rischi aventi probabilità e gravità diverse per i diritti e le libertà delle persone fisiche costituiti dal trattamento, sia al momento di determinare i mezzi del trattamento sia all'atto del trattamento stesso.

Nelle linee guida n. 4/2019 sull'articolo 25 del Regolamento il Comitato ha chiarito che il fulcro della disposizione è garantire una adeguata ed efficace protezione dei dati fin dalla progettazione e una protezione per impostazione predefinita, il che significa che i titolari dovrebbero essere in grado di dimostrare che incorporano nel trattamento le misure e le garanzie adeguate ad assicurare l'efficacia dei principi di protezione dei dati, dei diritti e delle libertà degli interessati ed ha invitato i titolari a tenere conto, nell'ambito della progettazione e impostazione del trattamento in un'ottica privacy oriented, anche degli obblighi di fornire una tutela specifica ai minori e ad altri gruppi di soggetti vulnerabili.

Le linee guida 2/2019 sul trattamento di dati personali ai sensi dell'articolo 6, paragrafo 1, lettera b), del Regolamento generale sulla protezione dei dati nel contesto della fornitura di servizi online agli interessati, al paragrafo 26, recitano che qualora un titolare non sia in grado di dimostrare a) l'esistenza di un contratto, b) la validità del contratto ai sensi del diritto contrattuale nazionale applicabile e c) l'oggettiva necessità del trattamento ai fini dell'esecuzione del contratto, tale titolare dovrebbe prendere in considerazione un'altra base giuridica per il trattamento.

Dall'istruttoria è emerso che, alla data del 30 marzo 2023, la Società non prevedeva alcun meccanismo per la verifica dell'età degli utenti all'atto della registrazione al servizio ChatGPT, ancorché le condizioni di servizio, nella versione a quel tempo vigente e aggiornata al 14 marzo 2023, individuassero i minori di età (compresa tra i 13 ed i 18 anni) tra i potenziali utenti, e stabilissero che, in tal caso, per perfezionare un valido vincolo contrattuale, fosse necessario il consenso del titolare della responsabilità genitoriale.

Prima dell’intervento dell’Autorità, quindi, tutti gli utenti, compresi i minori di età, potevano iscriversi al servizio ChatGPT ed utilizzarlo senza che venisse chiesto loro di sottoporsi ad alcuna verifica dell’età. Come già chiarito in sede di contestazione, ad avviso dell’Autorità, l’assenza di uno standard comune idoneo a garantire, in maniera certa e assoluta, l’efficacia di un modello di verifica dell’età dell’utente e la discussione tutt’ora in atto a livello europeo al riguardo, non possono essere considerate ragioni idonee ad escludere l’adempimento degli obblighi a cui è tenuto il titolare del trattamento, segnatamente quello di verificare l’effettiva capacità negoziale dell’utente ai fini della validità del contratto.

Nella memoria difensiva, il titolare ha, tra l’altro, rappresentato che i termini di utilizzo aggiornati al 14 marzo 2023 e disponibili online il 30 marzo 2023 prevedevano che Qualora l'utente avesse meno di 18 anni, [dovesse] avere il permesso dei genitori o del tutore legale per utilizzare i Servizi e che l'uso del termine "permesso" nella frase di cui sopra non si riferisce al consenso ai sensi dell'articolo 6, par. 1, lettera a), del GDPR. Si tratta di garantire che un adolescente confermi di avere il permesso dei genitori o del tutore legale per utilizzare ChatGPT, anche per sostenere la validità del contratto tra OpenAI e l'adolescente.

Tuttavia, da quanto risulta in atti, contrariamente a quanto prescritto dalle linee guida 2/2019 sopracitate, il titolare, alla data del 30 marzo 2023, non ha predisposto sistemi atti a verificare l’età dei soggetti per garantire che un minore di età compresa tra i 13 ed i 18 anni confermasse di avere il permesso dei genitori o del tutore legale per utilizzare ChatGPT, anche per sostenere la validità del contratto tra OpenAI e l’adolescente.

Con ogni evidenza, in assenza di tali sistemi, il titolare non era nelle condizioni di poter appurare che il perfezionamento del vincolo contrattuale avvenisse con l’effettivo coinvolgimento dei titolari della responsabilità genitoriale e, pertanto, si esclude che lo stesso fosse in grado di dimostrare la validità del contratto ai sensi del diritto contrattuale nazionale applicabile (cfr. le linee guida 2/2019).

Da quanto precede emerge che, alla data del 30 marzo 2023, il titolare non avesse, ex art. 24, messo in atto misure volte a garantire che il trattamento dei dati all’atto di iscrizione al servizio ChatGPT fosse conforme al Regolamento, né, ex art. 25, che la Società avesse adottato misure tecniche e organizzative volte ad attuare in modo efficace i principi di protezione dei dati, quali la minimizzazione, e a integrare nel trattamento le necessarie garanzie al fine di soddisfare i requisiti del presente regolamento e tutelare i diritti degli interessati.

Con riferimento a quanto contestato, il titolare ha dichiarato che Gli utenti, compresi quelli di età inferiore ai 18 anni, hanno avuto la possibilità di opporsi al trattamento delle loro conversazioni ChatGPT ai fini di addestramento fin dall'inizio, quando ChatGPT è stato rilasciato per la prima volta il 30 novembre 2022. Inizialmente, gli utenti potevano rivolgersi all'assistenza di OpenAI per opporsi al trattamento. Con la crescita del numero di utenti di ChatGPT e l'espansione delle operazioni di OpenAI, nel febbraio 2023 OpenAI ha creato un modulo online (Allegato 4 della lettera del 4 aprile 2023) per consentire agli utenti di opporsi all’utilizzo dei loro contenuti per addestrare e migliorare i modelli di OpenAI. Il titolare ha dichiarato altresì che Per gli utenti che non si sono opposti all’utilizzo dei loro dati per le finalità di addestramento, OpenAI ha adottato misure di protezione della privacy-by-design, come prescritto dall’articolo 25 GDPR, prima che le conversazioni potessero essere utilizzate nel processo di addestramento. In particolare, OpenAI ha dissociato le conversazioni dall'account dell'utente, ha utilizzato un processo di filtraggio per rimuovere gli identificatori personali che gli utenti potrebbero aver inserito nelle loro conversazioni e ha incorporato una fase di revisione umana per confermare che qualsiasi richiesta con informazioni identificabili o sensibili fosse esclusa dai dataset etichettati che possono essere utilizzati per l'addestramento del modello. I dati delle conversazioni possono essere utilizzati per la messa a punto per migliorare la funzionalità e la sicurezza del modello solo dopo le fasi di de-identificazione.

I chiarimenti in merito al diritto di opposizione dell'interessato, il cui esercizio è stato agevolato a seguito dell'intervento del Garante, risultano irrilevanti ai fini dell'accertamento della violazione de qua, in quanto attinenti ad un momento del trattamento successivo a quello oggetto di analisi.

Invero, quanto accertato dall'Autorità riguarda l'omessa previsione da parte della Società di verifiche funzionali ad impedire l'accesso al servizio ad interessati minori di 13 anni ed a garantire il coinvolgimento nel processo di iscrizione del titolare della responsabilità genitoriale per i minorenni di età compresa tra i 13 ed i 18 anni, ai fini dell'ottenimento della relativa autorizzazione, come richiesto dai termini e dalle condizioni contrattuali definiti dalla Società stessa.

L'Autorità, invece, non ritiene sussistano idonei elementi per ritenere accertata la violazione, contestata ex art. 166, par.5, del Codice, relativa al consenso digitale dei minori di cui all'art. 8 del Regolamento atteso che la base giuridica di riferimento, come sopra illustrato, è stata individuata nell'esecuzione del contratto ex art. 6, par. 1, lett. b), del Regolamento.

Per quanto sopra rappresentato, si ritiene che OpenAI abbia violato gli artt. 24 e 25, par. 1 del Regolamento e che detta violazione, accertata alla data del 30 marzo 2023, debba ritenersi consumata il 30 novembre 2022 e sia di natura non continuata.

6.5 ART. 83, PAR. 5, LETT. E), DEL REGOLAMENTO: CAMPAGNA INFORMATIVA

L'Ufficio ha contestato ad OpenAI la violazione dell'art. 83, par. 5, lett. d), del Regolamento per omessa osservanza di un ordine dell'Autorità, segnatamente l'ordine di messa in conformità di cui al punto 9 del provvedimento n. 114/2023.

L'art. 58, par. 2, del Regolamento nell'elencare i poteri correttivi di cui dispone ogni Autorità di controllo prevede, alla lett. d), che le stesse possano ingiungere al titolare del trattamento di conformare i trattamenti alle disposizioni del Regolamento, se del caso, in una determinata maniera ed entro un determinato termine.

L'art. 83, par. 5, lett. e), del Regolamento prevede che l'inosservanza di un ordine, di una limitazione provvisoria o definitiva di trattamento o di un ordine di sospensione dei flussi di dati dell'autorità di controllo ai sensi dell'art. 58, par. 2, costituisca una violazione soggetta alla sanzione amministrative pecuniaria fino a 20 milioni di euro o per le imprese, fino al 4 % del fatturato mondiale totale annuo dell'esercizio precedente.

Con il provvedimento n. 114/2023 il Garante ha prescritto ad OpenAI, come condizione per la sospensione dell'efficacia del provvedimento di limitazione provvisoria n.112/2023, la realizzazione, entro il 15 maggio 2023, di una campagna di informazione di natura non promozionale su tutti i principali mezzi di comunicazione di massa italiani (radio, televisione, giornali e Internet) da concordare con il Garante, allo scopo di informare le persone dell'avvenuta probabile raccolta dei loro dati personali ai fini dell'addestramento degli algoritmi, dell'avvenuta pubblicazione sul sito web della Società di un'apposita informativa di dettaglio e della messa a disposizione, sempre sul sito web della Società di uno strumento attraverso il quale tutti gli interessati avrebbero potuto chiedere e ottenere la cancellazione dei propri dati personali (punto 9 del provvedimento n. 114/2023).

Con nota del 15 maggio (prot. n. 78218/23) OpenAI ha comunicato di aver adempiuto alla prescrizione dell'Autorità attraverso le misure illustrate nella ricostruzione in fatto (cfr. par. 2), tra cui un video didattico da realizzare in collaborazione con l'Autorità. Con riferimento a quest'ultima iniziativa il titolare, con la nota citata, ha rappresentato al Garante di aver predisposto il video della campagna informativa, la volontà di adottarlo in collaborazione con l'Autorità, rendendolo pubblico tramite i relativi canali istituzionali.

L’Autorità, con nota del 18 maggio successivo (prot. n. 79806/23), ha espresso forte contrarietà rispetto alla campagna mediatica, in quanto realizzata senza alcuna preventiva informazione e nessun previo accordo con il Garante in merito ai termini ed alle condizioni della stessa, contrariamente a quanto prescritto nel provvedimento n. 114/2023. Con la stessa nota l’Autorità si è riservata ogni valutazione rispetto al puntuale adempimento delle prescrizioni impartite ed ha invitato la Società a presentare senza ritardo e, comunque, non oltre il successivo 19 maggio (prorogato al 23 maggio), un progetto di campagna di comunicazione in linea con la prescrizione e con le osservazioni rappresentate.

Con nota del 23 maggio 2023 (prot. n. 82299/23), la Società, in un’ottica di adeguamento alle indicazioni ricevute, ha rappresentato la volontà di volersi impegnare al fine di organizzare con emittenti televisive e radiofoniche nazionali la diffusione di messaggi informativi inerenti al servizio ChatGPT.

Con nota del 22 giugno 2023 (prot. n. 97898/23), l’Autorità, ha comunicato di non essere in grado, alla luce delle informazioni fornite, di valutare l’impatto della campagna con particolare riferimento alle iniziative che la Società aveva intenzione di lanciare via social, mancando utili elementi di valutazione ed ha giudicato le prospettate attività mediante giornali e televisioni al di sotto delle aspettative e difficilmente idonee a raggiungere il pubblico sperato. Il Garante ha espresso altresì parere contrario all’utilizzazione del proprio logo nella menzionata campagna comunicazionale.

Nella memoria difensiva (cfr. memoria difensiva, versione tradotta, pag. 59) ed in sede di audizione OpenAI ha sostenuto di aver correttamente adempiuto alla misura correttiva de qua considerate le iniziative poste in essere nonostante il ristretto lasso di tempo concesso (poco più di un mese), la necessità di allocare un budget e le ridotte risorse aziendali (anche in termini di esperienza specifica in ambito comunicativo) o, al limite, di aver semplicemente “mal interpretato il Provvedimento n. 114” e, dopo la comunicazione dell’Autorità del 18 maggio 2023, di aver “lavorato duramente per superare i suoi limiti interni” e soddisfare le aspettative del Garante.

La Società, infine, ha dichiarato di volersi impegnare “a collaborare con il Garante e adottare i necessari sforzi per aumentare la trasparenza e sensibilizzare gli interessati in generale in merito ai loro diritti e allo sviluppo di sistemi di IA”.

L’Autorità ritiene che OpenAI non abbia adempiuto a quanto dalla stessa prescritto, ai sensi dell’art. 58, par. 2, lett. d), del Regolamento, con riferimento al punto 9 del provvedimento n. 114/2023, in quanto la campagna informativa promossa, come realizzata alla data del 15 maggio 2023, non è stata concordata con l’Autorità, né la stessa è risultata idonea, per la scelta dei mezzi e delle modalità di comunicazione, nonché per il tempo assai limitato della stessa, a raggiungere la generalità del pubblico interessato dai servizi ChatGPT, tra l’altro, rendendola edotta dei diritti riconosciutile dalla disciplina vigente a cominciare dal diritto ad opporsi ai trattamenti di dati personali posti in essere dalla Società tra l’altro per l’addestramento dei propri modelli e, ciò, con particolare riferimento, ai dati personali dei non utenti, non destinatari di qualsivoglia diversa forma di informazione.

Per quanto sopra rappresentato, si ritiene che OpenAI abbia violato l’art. 83, par. 5, lett. e), del Regolamento per inosservanza di un ordine dell’Autorità ex art. 58, par. 2, lett. d), del Regolamento e che detta violazione debba ritenersi consumata il 15 maggio 2023 e sia di natura non continuata.

6.6 ART. 5, PAR. 1, LETT. D), DEL REGOLAMENTO

L’Ufficio ha contestato ad OpenAI la violazione dell’art. 5, par. 1, lett. d), del Regolamento in quanto il servizio ChatGPT, alla data del 30 marzo 2023, generava output inesatti.

L’art. 5, par. 1, lett. d) del Regolamento prevede che i dati personali debbano essere esatti e, se necessario, aggiornati (principio di esattezza) e che i titolari debbano adottare tutte le misure ragionevoli per cancellare o rettificare tempestivamente i dati inesatti rispetto alle finalità per le quali sono trattati.

Con il provvedimento n.112/23 di limitazione del trattamento, adottato d’urgenza il 30 marzo 2023, il Garante ha messo in evidenza il fatto che le informazioni fornite da ChatGPT negli output agli utenti non sempre corrispondevano al dato reale.

La questione tecnico-giuridica relativa all’accuratezza dei LLM è uno degli argomenti maggiormente dibattuti in dottrina ed è stata oggetto di valutazione anche da parte della task force europea creata dall’EDPB il 13 aprile 2023. Nel report finale, approvato in data 23 maggio 2024 e pubblicato il successivo 24 maggio, la task force ha rilevato come lo scopo del trattamento dei dati di OpenAI sia quello di addestrare il modello linguistico GPT sotteso ai servizi ChatGPT e non necessariamente quello di fornire informazioni accurate, in quanto la natura probabilistica del sistema porta il modello a produrre risultati parziali o discriminatori (biased). Tuttavia, è probabile che i risultati forniti da ChatGPT siano considerati di fatto accurati dagli utenti finali indipendentemente dalla loro effettiva accuratezza. Risulta, pertanto, importante che il titolare del trattamento fornisca informazioni adeguate sui meccanismi probabilistici di creazione degli output e sul loro limitato livello di affidabilità, compreso un riferimento esplicito al fatto che il testo generato, sebbene sintatticamente corretto, possa essere distorto o discriminatorio. Tale adempimento, riconducibile al principio di trasparenza di cui all’art. 5, par. 1, lett. a), del Regolamento è utile ad evitare un’interpretazione errata degli output di ChatGPT da parte degli utenti, ma non esime la Società dal doversi adoperare per il rispetto del principio di esattezza (cfr. final report, pagg. 8 e 9). A tal riguardo, si osserva come OpenAI, a seguito del provvedimento del Garante n. 114/2023, pur dando atto delle difficoltà tecniche, abbia implementato alcune misure per ridurre gli effetti dell’inaccuratezza degli output. Nella memoria difensiva la Società ha sottolineato come sin dal lancio del servizio nel novembre 2022, OpenAI ha chiarito agli utenti che ChatGPT non deve essere inteso come una fonte accurata di informazioniattraverso dichiarazioni, avvisi, articoli, FAQ sul sito web ed una finestra pop-up dedicata per informare gli utenti sulla potenziale imprecisione delle risposte del servizio (cfr. memoria difensiva, versione tradotta, pagg. 45 e 51-52). La Società ha rappresentato quali misure sono state adottate, in ogni fase di addestramento, per identificare e rimuovere informazioni imprecise, inaffidabili o potenzialmente dannose (fase di pre-addestramento), istruire i modelli a rifiutare di fornire informazioni private o sensibili relative a persone fisiche (fase di post-addestramento) e offrire alle persone di segnalare inesattezze e chiederne la rettifica (fase di utilizzo del servizio). La Società ha, inoltre, precisato che, sebbene non sia tecnicamente possibile, rendere un LLM accurato al 100%, “OpenAI si impegna a migliorare l’accuratezza dei suoi modelli, non perché si debba fare affidamento su ChjatGPT come fonte accurata di informazioni, ma perché risposte più accurate renderanno ChatGPT più utile ai suoi utenti (cfr. memoria difensiva, versione tradotta, pag. 47) e che uno studio del novembre 2023 ha riconosciuto a ChatGPT (modelli GPT 4, GPT 4 turbo e GPT 3.5) il più basso tasso di allucinazioni tra i principali servizi di intelligenza artificiale generativa (cfr. memoria difensiva, versione tradotta, pagg. 47-48).

Alla luce di quanto sopra esposto pare evidente che la questione giuridica relativa all’inesattezza dei dati personali degli output del servizio ChatGPT sollevata dal Garante nel provvedimento d’urgenza di limitazione provvisoria n. 112/2023 sia tutt’altro che risolta e che le misure tecniche ed organizzative implementate da OpenAI nel corso dell’ultimo anno siano frutto di un processo costante ed in fieri. Come riferito dalla Società i modelli e i meccanismi di accuratezza di OpenAI sono, per loro natura, in continuo sviluppo (cfr. memoria difensiva, versione tradotta, pag. 45). La violazione contestata deve, dunque, qualificarsi come una violazione continuativa (ovverosia di natura permanente) e quindi ancora in essere alla data 15 febbraio 2024, dies a quo dello stabilimento della Società nell’Unione europea, attraverso la società consociata OpenAI Ireland

Ltd. Dovendosi applicare da tale data il meccanismo dello sportello unico con conseguente passaggio della competenza all'autorità di controllo capofila, ai sensi dell'art. 56 del Regolamento (individuata nella Data Protection Commission irlandese), sulla scorta del parere dell'EDPB n. 8/2020 sulla competenza di un'autorità di controllo in caso di mutamento delle circostanze relative allo stabilimento principale o unico, si ritiene di non poter, per questo specifico profilo, procedere per difetto di competenza e si dispone la trasmissione degli atti all'autorità capofila irlandese.

7. ORDINANZA INGIUNZIONE PER L'APPLICAZIONE DELLA SANZIONE AMMINISTRATIVA PECUNIARIA E DELLE SANZIONI ACCESSORIE

L'Autorità, ai sensi dell'art. 58, par. 2, lett. i), e 83 del Regolamento nonché dell'art. 166 del Codice, ha il potere di infliggere una sanzione amministrativa pecuniaria ai sensi dell'art. 83, in aggiunta o in luogo delle altre misure correttive previste nel medesimo paragrafo.

Nel caso di specie, atteso che le misure correttive ai sensi dell'art. 58, par. 2, lett. d), del Regolamento sono già state disposte dall'Autorità con il provvedimento di sospensione del provvedimento urgente di limitazione provvisoria n. 114/2023, con il presente provvedimento il Garante adotta l'ordinanza ingiunzione con la quale dispone l'applicazione della sanzione amministrativa pecuniaria ai sensi dell'art. 58, par. 2, lett. i), del Regolamento e la sanzione accessoria di realizzare una campagna di comunicazione istituzionale volta a promuovere la consapevolezza del diritto alla protezione dei dati personali, ai sensi dell'art. 166, comma 7, del Codice, con particolare riferimento al diritto degli interessati di esercitare il diritto di opposizione.

Nella determinazione della sanzione l'Autorità tiene conto dei principi e dell'interpretazione in materia fornita dall'EDPB nelle linee guida 4/2022 sul calcolo delle sanzioni amministrative pecuniarie ai sensi del GDPR, versione 2.1, adottate il 24 maggio 2023.

Sulla scorta delle argomentazioni sopra addotte, il Garante ha accertato la violazione delle seguenti disposizioni del Regolamento: art. 33; artt. 5, par. 2 e 6; artt. 5, par. 1, lett. a), 12 e 13; artt. 24 e 25, par. 1; art. 83, par. 5, lett. e), del Regolamento.

Nel caso di specie, occorre innanzitutto rilevare che la Società ha posto in essere una serie di condotte che hanno integrato più violazioni, come nei paragrafi precedenti specificamente delineate e motivate. Le violazioni relative alla base giuridica (artt. 5, apr par. 2 e 5 del Regolamento), alla trasparenza (artt. 5, par. 1, lett. a), 12 e 13) e all'age verification (artt. 24 e 25, par. 1) possono essere ricondotte, per il principio di unità dell'azione, sotto l'egida dell'art. 83, par. 3, del Regolamento, secondo il quale in presenza di più violazioni del Regolamento, relative allo stesso trattamento o a trattamenti collegati, l'importo totale della sanzione amministrativa pecuniaria non può superare l'importo previsto per la violazione più grave. Segnatamente, con riferimento a tali violazioni, è configurabile un'ipotesi di trattamenti collegati, come definita al paragrafo 28 delle citate linee guida (una condotta unitaria consiste in più azioni che sono poste in essere sulla base di una volontà unitaria e sono contestualmente, spazialmente e temporalmente correlate in modo così stretto da potere essere considerate, da un punto di vista oggettivo, come un'unica condotta coerente). La violazione più grave tra queste deve essere ravvisata nella violazione degli obblighi di trasparenza atteso che sia l'art. 5, par. 1, lett. a) (principio di trasparenza) che gli artt. 12 e 13 (diritti degli interessati) sono sanzionati ai sensi dell'art. 83, par. 5, il quale fissa il massimo edittale nella somma di 20 milioni di euro ovvero, per le imprese, nel 4% del fatturato mondiale annuo dell'esercizio precedente ove superiore. Per contro, le violazioni relative al data breach (art. 33 del Regolamento) e all'inosservanza di un ordine dell'Autorità (art. 83, par. 5, lett. e), del Regolamento) costituiscono violazioni separate, non rientranti nell'ambito di applicazione dell'art. 83, par. 3, del Regolamento.

Ai sensi dell'art. 83, par. 1 del Regolamento, la sanzione amministrativa deve essere effettiva, proporzionata e dissuasiva in relazione al singolo caso. Nelle citate linee guida l'EDPB ha

precisato che il calcolo delle sanzioni amministrative pecuniarie debba iniziare da un punto di partenza armonizzato, che costituisce la base iniziale per l'ulteriore calcolo dell'ammontare della sanzione, in cui sono prese in considerazione e ponderate tutte le circostanze del caso (cfr. par. 46). Il punto di partenza armonizzato deve prendere in considerazione tre fattori: 1) natura della violazione ai sensi dell'art. 83, parr. da 4 a 6, del Regolamento; 2) gravità della violazione; 3) fatturato dell'impresa (cfr. par. 48).

Partendo dal primo profilo, nel caso di specie, si riscontrano due violazioni, in astratto, di natura più grave (art. 83, par. 5, del Regolamento) ed una violazione meno grave (art. 83, par. 4, del Regolamento). Le prime due si riferiscono alle tre violazioni, considerate in maniera unitaria ex art. 83, par. 3, del Regolamento e all'inosservanza dell'ordine del Garante, mentre la seconda concerne il data breach.

Quanto alla gravità in concreto, gli elementi da prendere in considerazione sono: a) natura, gravità e durata della violazione (art. 83, par. 2, lett. a), del Regolamento); b) carattere doloso o colposo della violazione (art. 83, par. 2, lett. b), del Regolamento); c) categorie di dati personali interessate dalla violazione (art. 83, par. 2, lett. g), del Regolamento).

Nella fattispecie in esame, la gravità delle violazioni, con riferimento alle tre violazioni collegate dal principio di unità dell'azione, deve essere considerata di livello elevato atteso che: i) la natura delle violazioni attiene a due principi fondamentali (trasparenza e accountability, segnatamente, all'incapacità del titolare di dimostrare che l'individuazione e l'elezione della base giuridica del trattamento sia avvenuta prima che il trattamento avesse inizio) ed ai diritti degli interessati (con riferimento alle violazioni degli obblighi informativi, segnatamente le informazioni che avrebbero dovuto essere rese in merito alle attività di trattamento per fini di addestramento dei modelli sottesi al servizio ChatGPT, in particolare in relazione ai non utenti, le cui aspettative in merito al trattamento dei dati per le suddette finalità sono da considerare praticamente inesistenti); ii) la natura del trattamento comporta rischi significativamente elevati in quanto connesso ad una tecnologia innovativa, dirompente ed in rapidissima espansione; iii) l'oggetto del trattamento ha natura transfrontaliera di portata globale con effetti praticamente incontrollabili da parte dei soggetti interessati; iv) la finalità del trattamento rientra nel core business della Società; v) il numero di interessati convolti è altissimo: 1,8 milioni di utenti italiani mensili attivi del servizio ChatGPT nel marzo 2023, ma soprattutto potenzialmente l'intera popolazione italiana a cui sia riconducibile una informazione pubblicamente disponibile sul web raccolta, direttamente o indirettamente da OpenAI per l'addestramento dei modelli GPT sottesi al servizio ChatGPT; vi) la natura dei dati ha verosimilmente riguardato anche dati particolari e, in assenza di meccanismi di verifica dell'età e di sistemi di filtraggio dei dati di minori ai fini di addestramento, informazioni personali riferite a soggetti minorenni. La limitata durata della violazione (dal 30 novembre 2022 al 30 marzo 2023) non pare costituire elemento idoneo a controbilanciare il giudizio di alta gravità in quanto la fine della violazione è dipesa e coincisa con l'intervento d'urgenza del Garante.

Per contro, con riferimento alla violazione dell'art. 33 del Regolamento, sono elementi idonei a configurare di livello basso la relativa violazione il limitato numero di utenti potenzialmente coinvolti nel data breach (440 utenti italiani potenziali), la natura dell'incidente di sicurezza (bug in una libreria open source), la natura comune dei dati compromessi, il rapido ed efficace intervento della Società nel mettere in sicurezza (sospendendo il servizio), rimediare a livello aziendale (fissando il bug) e a livello di comunità (intervento sulla libreria open source), nonché l'avvenuta notifica del data breach, sebbene ad una autorità di controllo pacificamente non competente all'epoca dei fatti.

Di contro, deve ritenersi di livello elevato la violazione dell'art. 83, par. 5, lett. e), del Regolamento attesa la gravità della stessa , tenuto conto della peculiarità dell'ordine dell'Autorità cui era subordinata la riattivazione del servizio in Italia secondo quanto disposto dal punto 9 del provvedimento n. 114/2023, non rilevando le difficoltà riscontrate nell'adempimento a causa delle

dimensioni e dell'organizzazione della Società all'epoca dei fatti, del tempo limitato per l'adempimento.

Tutte le violazioni devono ritenersi di natura colposa. Come affermato dal Gruppo di lavoro Art. 29, nelle linee guida riguardanti l'applicazione e la previsione delle sanzioni amministrative pecuniarie ai fini del regolamento (UE) n. 2016/679, adottate il 3 ottobre 2017 e recepite dall'EDPB il 25 maggio 2018 (linee guida WP 253), il comportamento doloso si riferisce sia alla consapevolezza che all'intenzionalità (coscienza e volontà) di commettere un illecito, mentre nel comportamento colposo manca l'intenzione di causare la violazione nonostante il mancato rispetto di un obbligo di diligenza. OpenAI, richiamando una recente pronuncia della Corte di Giustizia dell'Unione europea (sentenza C-807/21 del 5 dicembre 2023), secondo cui è onere dell'autorità di controllo stabilire che una violazione sia stata commessa con dolo o colpa dal titolare del trattamento in quanto solo le violazioni illecite possono comportare l'imposizione di una sanzione amministrativa pecuniaria, ha sostenuto che il Garante non abbia, nell'atto di contestazione delle violazioni ex art. 166 del Codice, fornito alcun elemento in proposito. La Società ha richiamato, al fine di argomentare la propria tesi difensiva, il fatto che OpenAI, società di ricerca statunitense senza scopo di lucro, non avesse pianificato o previsto di gestire un importante servizio rivolto ai consumatori, ma una volta che la base di utenti e le operazioni di ChatGPT sono cresciute tanto da rientrare nel campo di applicazione del GDPR, OpenAI si è rapidamente attivata per migliorare le sue pratiche. Di conseguenza, il carattere doloso o colposo delle presunte violazioni del GDPR è inesistente e non può essere presunto da Garante (cfr. memoria difensiva, versione tradotta, pag. 61-62). A tal proposito si osserva che, se è vero che la CGUE ha sancito nella menzionata decisione che l'art. 83 del Regolamento non consente di infliggere una sanzione amministrativa pecuniaria senza l'accertamento che tale violazione sia stata commessa con dolo o colpa dal titolare del trattamento (cfr. par. 75), è altresì vero che la stessa Corte ha fatto salvo il basilare principio ignorantia legis non excusat, affermando che un titolare del trattamento può essere sanzionato per un comportamento ricadente nell'ambito di applicazione del RGPD qualora detto titolare non potesse ignorare il carattere illecito del proprio comportamento, a prescindere dalla sua consapevolezza di violare le disposizioni del RGPD (cfr. par. 76). Tale principio era già stato enunciato dalla Corte di Giustizia in altro caso (sentenza C-601/16 del 25 marzo 2021, par. 97 e 98) in cui aveva sostenuto che un'impresa può essere sanzionata per un comportamento rientrante nell'ambito di applicazione dell'articolo 101, paragrafo 1, TFUE qualora tale impresa non potesse ignorare il carattere anticoncorrenziale del suo comportamento, indipendentemente dal fatto che fosse o meno consapevole di violare le regole di concorrenza del Trattato (v., in tal senso, sentenza del 18 giugno 2013, Schenker & Co. e a., C 681/11, EU:C:2013:404, punto 37). Ne consegue che il fatto che tale impresa abbia qualificato erroneamente in diritto il suo comportamento su cui si fonda l'accertamento dell'infrazione non può avere l'effetto di esonerarla dall'irrogazione di un'ammenda, in quanto essa non poteva ignorare il carattere anticoncorrenziale di tale comportamento (sentenza del 18 giugno 2013, Schenker & Co. e a., C 681/11, EU:C:2013:404, punto 38)". Nel caso di specie, OpenAI non può sottrarsi all'obbligo di conoscenza e di applicazione del Regolamento semplicemente sostenendo che non poteva prevedere il successo del suo servizio. La conoscenza e l'applicazione del Regolamento prescindono dal successo di una iniziativa economica in quanto il Regolamento, come noto, tutela un diritto fondamentale previsto e protetto dall'art. 8 della Carta dei diritti fondamentali dell'Unione europea e deve essere rispettato prima di effettuare un trattamento di dati personali, non in un momento successivo, discrezionalmente individuato dal titolare. Il fatto stesso che la Società abbia considerato di dover rispettare il Regolamento solo a fronte del successo riscosso dal servizio ChatGPT nell'Unione europea integra di per sé la negligenza sottesa al concetto di colpa e consente di ritenere dimostrata la sussistenza di tale elemento soggettivo in capo alla Società. Più nello specifico, la colpa è dimostrata proprio dalla consapevolezza manifestata dalla Società di aver adottato sin dall'inizio … un modello di compliance privacy basato principalmente su tre elementi: la trasparenza, la privacy by design/il principio di minimizzazione e meccanismi di opt-out per gli utenti e che alla luce del grande successo di ChatGPT, sono state intraprese continue attività di compliance tese al

miglioramento del modello organizzativo (a titolo esemplificativo, bozze DPIA e LIA). (cfr. verbale audizione pag. 3). Non può, infine, trovare accoglimento l'argomento difensivo secondo cui OpenAI era (ed è tuttora) una organizzazione di ricerca di piccole dimensioni con risorse limitate… (cfr. documento allegato al verbale di audizione, pag. 6), atteso che, come riferito nelle sopra citate linee guida WP 253, Le imprese dovrebbero essere responsabili dell'adozione di strutture e risorse idonee alla natura e alla complessità della propria attività. Pertanto, i titolari del trattamento e i responsabili del trattamento non possono legittimare violazioni della normativa sulla protezione dei dati appellandosi a una carenza di risorse. La colpa in capo ad OpenAI deve pertanto essere considerata grave in quanto la Società ha lanciato un servizio che comporta il trattamento su larga scala di dati personali a livello mondiale, senza aver adottato le misure minime di compliance con il Regolamento.

Sempre ai fini della quantificazione della sanzione pecuniaria amministrativa rilevano i fattori aggravanti di cui all'art. 83, par. 2, lett. d) e k) del Regolamento.

Quanto al primo profilo, il grado di responsabilità del titolare del trattamento deve ritenersi elevato a causa della mancata adozione, al momento di lancio del servizio, di misure tecniche ed organizzative idonee a mitigare i rischi per i diritti e le libertà degli interessati ed attribuire loro l'esercizio delle prerogative di cui al Capo III del Regolamento. Quanto alla seconda circostanza, si rileva che le violazioni accertate hanno consentito alla società di avvalersi di un vantaggio competitivo e di conseguenza di ottenere benefici finanziari.

Ai fini della adozione della sanzione amministrativa, si tiene conto a titolo di fattore attenuante l'adozione delle misure poste in essere dal titolare del trattamento per porre rimedio alla violazione e attenuarne i possibili effetti negativi (art. 83, par. 2, lett. f, del Regolamento), in particolare:

- l'aggiornamento dell'informativa di cui agli articoli 12 e 13 del Regolamento, anche con riferimento alle finalità di training dei modelli, in ultimo con l'informativa privacy del 15 dicembre 2023, in vigore dal 15 febbraio 2024;

- la pubblicazione della privacy policy nella home page del sito, nelle sezioni privacy policy, nell'help center del sito e nella pagina di log-in ovvero la sua collocazione nella pagina di registrazione in una posizione tale da consentirne la lettura prima della conclusione della procedura di registrazione, nonché la presentazione agli utenti già registrati di una finestra che riportava i link alla privacy policy ed all'help center unitamente alla richiesta della conferma della loro età;

- l'implementazione dei meccanismi di age gate descritti nel par. 3.1.4, sia per i nuovi utenti che per gli utenti già registrati;

- con riferimento all'addestramento dei modelli, le misure adottate per limitare il trattamento dei dati personali durante le varie fasi di addestramento.

Il fatturato di OpenAI rilevante per il calcolo della sanzione è il fatturato totale mondiale riferito all'esercizio dell'anno 2023. I dati fiscali forniti dalla Società, con nota del 23 aprile 2024, indicano tale valore in 1.029.186.389,00 dollari, pari a circa 948.487.890,00 euro alla data di adozione del presente provvedimento.

In ragione dei suddetti elementi, valutati nel loro complesso, tenuto conto del fatturato mondiale totale annuo dell'esercizio precedente della Società, si ritiene di determinare, l'importo complessivo della sanzione amministrativa pecuniaria in euro 15.000.000,00, pari a circa all'1,58% del fatturato mondiale totale annuo dell'esercizio 2023. Tale importo risulta determinato nei termini che seguono:

- ai sensi dell'art. 83, par. 3, del Regolamento, ritenuta l'unicità della condotta trattandosi di

trattamenti collegati per le ragioni sopra addotte, l'ammontare della sanzione pecuniaria per la violazione degli artt. 5, par. 1, lett. a), 5, par. 2, 6, 12, 13, 24 e 25, par. 1, del Regolamento, viene calcolato nella in misura di euro 9.000.000,00;

• l'ammontare della sanzione pecuniaria per la violazione dell'art. 33 del Regolamento viene calcolato in misura pari a euro 320.000,00;

• l'ammontare della sanzione pecuniaria per la violazione dell'art. 83, par. 5, lett. e), del Regolamento viene calcolato in misura pari a euro 5.680.000,00.

La sanzione complessiva totale della sanzione amministrativa pecuniaria determinata in euro 15.000.000,00 viene considerata, ai sensi dell'art. 83, par. 1, del Regolamento, effettiva, proporzionata e dissuasiva.

L'interesse generale rispetto alla tematica dell'impatto dell'intelligenza artificiale generativa sul diritto alla protezione dei dati personali e la circostanza che la campagna di comunicazione che si era a suo tempo prescritto a OpenAI di porre in essere ma che la società non ha adeguatamente realizzato avrebbe dovuto valere, in particolare, a rendere edotti gli interessati dei diritti loro spettanti e delle modalità concrete attraverso le quali esercitarli in maniera effettiva impone, inoltre, l'applicazione della sanzione accessoria prevista dall'art. 166, comma 7, del Codice dell'ingiunzione a realizzare una campagna di comunicazione istituzionale volta alla promozione della consapevolezza delle questioni di protezione dei dati personali che vengono in rilievo nella fattispecie oggetto del presente procedimento con particolare riferimento proprio a termini e modalità attraverso i quali gli interessati potranno esercitare in maniera semplice e effettiva, tutti i diritti loro spettanti ai sensi della vigente disciplina a cominciare da quelli di opposizione, rettifica e cancellazione.

Tale sanzione è idonea a perseguire lo scopo di conoscibilità delle decisioni dell'Autorità sottesa al regime di pubblicità previsto dal legislatore alla luce del fatto che la campagna di informazione ordinata dal Garante nel provvedimento 114/2023, come sopra illustrato, non è stata correttamente adempiuta dalla Società in tal modo mancando la finalità perseguita dall'Autorità, attraverso tale ordine, di promuovere la consapevolezza e la comprensione del pubblico riguardo ai rischi connessi al servizio ChatGPT ed alle garanzie ed ai diritti in relazione ai trattamenti di dati personali effettuai nell'ambito di tale servizio.

Alla luce della gravità delle violazioni accertate, segnatamente tenuto conto che si tratta di un trattamento di larga scala che coinvolge un elevato numero di interessati, della natura delle violazioni e dei rischi in termini di protezione dei dati personali connessi alla messa a disposizione del pubblico di un servizio basato su una tecnologia innovativa e complessa in assenza delle dovute salvaguardie, si ritiene di ingiungere alla Società di realizzare una campagna di comunicazione istituzionale, da effettuare su tutti i principali mezzi di comunicazione italiani (radio, televisione, giornali e Internet), della durata di sei mesi decorrenti dall'avvio della campagna stessa da iniziare entro 45 giorni dalla notifica dell'approvazione da parte del Garante del piano di comunicazione,– che la Società dovrà fare pervenire entro 60 giorni dalla notifica del presente provvedimento ed- , i cui contenuti dovranno essere previamente approvati dal Garante stesso. La campagna dovrà essere finalizzata a promuovere la comprensione e la consapevolezza del pubblico in merito al funzionamento del servizio ChatGPT, alle implicazioni dello stesso rispetto al diritto alla protezione dei dati personali, con particolare riferimento alla raccolta dei dati di utenti e non-utenti per finalità di addestramento dei modelli ed ai diritti esercitabili, con specifico riferimento al diritto degli interessati di esercitare il diritto di opposizione e quello alla cancellazione dei dati.. Entro 60 giorni dal termine della campagna di comunicazione, inoltre, la società dovrà comunicare all'Autorità ogni informazione utile a valutare l'adempimento del presente ordine ivi incluse quelle relative alla modalità attraverso le quali agli interessati che ne abbiano fatto istanza siano stati o saranno garantiti i diritti di cui alla vigente disciplina ivi inclusi il diritto di opposizione e quello alla

cancellazione dei dati personali.

In caso di omesso tempestivo adempimento l'Autorità si riserva l'adozione di ulteriori provvedimenti.

Si ritiene, infine, che ricorrano i presupposti di cui all'art. 17 del Regolamento n. 1/2019 concernente le procedure interne aventi rilevanza esterna, finalizzate allo svolgimento dei compiti e all'esercizio dei poteri demandati al Garante, per l'annotazione delle violazioni qui rilevate nel registro interno dell'Autorità, previsto dall'art. 57, par. 1, lett. u) del Regolamento.

TUTTO CIÒ PREMESSO IL GARANTE

ai sensi dell'art. 57, par. 1, lett. f), del Regolamento, dichiara illecito il trattamento descritto nei termini di cui in motivazione effettuato da OpenAI OpCo LLC, con sede in 3180 18th Street, San Francisco, California, Stati Uniti d'America e conseguentemente:

ORDINA

a OpenAI OpCo LLC, con sede in 3180 18th Street, San Francisco, California, Stati Uniti d'America di pagare la somma complessiva di euro 15.000.000,00 a titolo di sanzione amministrativa pecuniaria per le violazioni degli artt. 5, par. 1, lett. a) e par. 2, 6, 12, 13, 24, 25, e 32 del Regolamento, rappresentando che il contravventore, ai sensi dell'art. 166, comma 8, del Codice ha facoltà di definire la controversia con il pagamento, entro il termine di sessanta giorni, di un importo pari alla metà della sanzione irrogata.

INGIUNGE

a) alla predetta Società, in caso di mancata definizione della controversia ai sensi dell'art. 166, comma 8, del Codice, di pagare la somma di euro 15.000.000,00, secondo le modalità indicate in allegato, entro 30 giorni dalla notificazione del presente provvedimento, pena l'adozione dei conseguenti atti esecutivi a norma dall'art. 27 della legge n. 689/1981;

b) quale sanzione accessoria, ai sensi dell'art. 166, par. 7, del Codice, per le ragioni espresse in motivazione, alla predetta Società di realizzare una campagna di comunicazione istituzionale, da effettuare su tutti i principali mezzi di comunicazione italiani (radio, televisione, giornali e Internet), della durata di sei mesi decorrenti dall'avvio della campagna stessa da iniziare entro 45 giorni dalla notifica dell'approvazione da parte del Garante del piano di comunicazione che la società dovrà fare pervenire entro 60 giorni dalla notifica del presente provvedimento - ed i cui contenuti dovranno essere previamente approvati dal Garante stesso. La campagna dovrà essere finalizzata a promuovere la comprensione e la consapevolezza del pubblico in merito al funzionamento del servizio ChatGPT, alle implicazioni dello stesso rispetto al diritto alla protezione dei dati personali, con riferimento alla raccolta dei dati di utenti e non-utenti per finalità di addestramento dei modelli ed ai diritti dagli stessi esercitabili ai sensi del Regolamento, con specifico riferimento al diritto degli interessati di esercitare il diritto di opposizione e quello di cancellazione dei dati personali;

c) alla Società di trasmettere all'Autorità entro sessanta giorni dal termine della campagna di comunicazione ogni informazione utile a valutare il corretto adempimento all'ordine di cui alla lettera che precede ivi inclusi i termini e le modalità attraverso i quali è stato e sarà garantito agli interessati l'esercizio dei diritti oggetto della campagna medesima a cominciare da quelli di opposizione e cancellazione.

DISPONE

a) ai sensi dell'art. 17 del Regolamento del Garante n. 1/2019, l'annotazione nel registro

interno dell'Autorità, previsto dall'art. 57, par. 1, lett. u) del Regolamento, delle violazioni e delle misure adottate;

b) la trasmissione degli atti all'autorità di controllo irlandese, in qualità di autorità di controllo capofila ai sensi dell'art. 56, par. 1, del Regolamento a far data dal 15 febbraio 2024, con riferimento alla violazione di cui all'art. 5, par. 1, lett. d) del Regolamento ed in relazione ad ulteriori profili di illiceità di trattamento dati di natura continuativa.

Ai sensi dell'art. 78 del Regolamento, nonché degli artt. 152 del Codice e 10 del d.lgs. 1° settembre 2011, n. 150, avverso il presente provvedimento può essere proposta opposizione all'autorità giudiziaria ordinaria, con ricorso depositato al tribunale ordinario del luogo ove ha la residenza il titolare del trattamento dei dati personali, o, in alternativa, al tribunale del luogo di residenza dell'interessato, entro il termine di trenta giorni dalla data di comunicazione del provvedimento stesso, ovvero di sessanta giorni se il ricorrente risiede all'estero.

Roma, 2 novembre 2024

IL PRESIDENTE
Stanzione

IL RELATORE
Cerrina Feroni

IL SEGRETARIO GENERALE
Mattei