UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Università degli studi di Padova

Dipartimento di Diritto Pubblico, Internazionale e Comunitario

Corso di Laurea in Diritto e Tecnologia

Tesi di Laurea

L'anonimizzazione dei dati personali. Tecniche e criticità.

Relatore
Prof.ssa Claudia Sandei

Laureando
Nicolò Falghera
n° matricola 2037659

Anno Accademico 2023/2024

Indice

Introduzione

1.Importanza della protezione dei dati personali e scopo della tesi

Nel contesto odierno, caratterizzato da una sempre più crescente digitalizzazione e dalla pervasività dei dati, la protezione delle informazioni personali è divenuta una priorità. In ogni istante, enormi quantità di dati vengono raccolte, trasmesse, analizzate e utilizzate (in sintesi, *trattate*) per numerosissime finalità: dalla conclusione dei più semplici contratti, al marketing; dalla ricerca scientifica alla sanità, da parte di persone, aziende e Pubblica Amministrazione.

A tal proposito, l'anonimizzazione dei dati emerge come una pratica cruciale per garantire il bilanciamento tra la necessità di utilizzo dei dati e la protezione della privacy degli individui. L'anonimizzazione, mediante diverse tecniche e metodologie, garantisce che i dati personali vengano resi non più riconducibili o associabili ad alcun individuo.

Nonostante le numerose criticità sollevate nel corso degli ultimi anni in merito al massivo utilizzo di dati personali, appare non particolarmente diffusa o consolidata la consapevolezza dell'importanza della protezione di questi. La presente tesi si propone, quindi, di analizzare la rilevanza dei metodi di anonimizzazione dei dati, mediante una valutazione delle diverse tecniche ad oggi utilizzate, dei loro vantaggi e delle loro criticità, mirando a fornire una panoramica completa delle sfide e delle opportunità legate alla protezione dei dati personali.

Iniziando da un'attenta revisione della (ad oggi) più rilevante normativa in ambito di protezione dei dati personali (Reg. UE 679/2016 – GDPR) e degli standard internazionali ISO/IEC 27559:2022, la tesi esplorerà le implicazioni pratiche dell'anonimizzazione dei dati personali, anche attraverso l'analisi di casi studio e l'esame delle pronunce dell'Autorità Garante per la Protezione dei Dati Personali, nonché le nuove frontiere della tecnologia.

2.Metodologia di ottenimento delle fonti

Al fine di garantire accuratezza e rilevanza dei temi trattati, il processo di ottenimento delle fonti è caratterizzato da un'attenta analisi di articoli scientifici (noti come *papers*) provenienti da istituti di ricerca e Università, nonché di atti e pubblicazioni ufficiali presenti su portali autorevoli di enti regolatori come il Garante per la Protezione dei Dati Personali e la Corte Suprema di Cassazione.

Il reperimento dei suddetti articoli è stato possibile principalmente grazie a database come Google Scholar.

L'impiego di *keywords* pertinenti e specifiche per il campo di studio ha facilitato l'individuazione dei documenti, così come l'utilizzo di filtri nella ricerca di informazioni aggiornate e rilevanti.

La quasi totalità delle fonti utilizzate è in lingua inglese, ed è stata mia premura tradurre gli incisi spesso riportati a piè di pagina con la massima precisione possibile, ai fini di preservarne il significato.

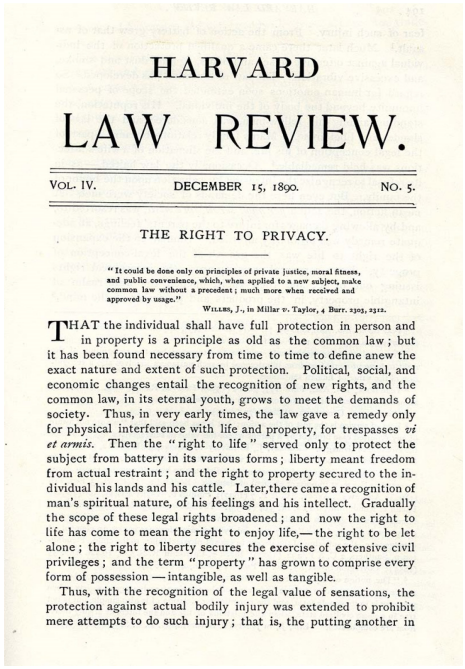
Capitolo 1: Normative sulla Protezione dei Dati Personali

1.Il Regolamento Generale sulla Protezione dei Dati (GDPR)

Da “The Right to Privacy” al GDPR

È doveroso iniziare effettuando una precisazione: i termini “privacy” e “protezione dei dati”, seppur spesso sovrapposti nel linguaggio ordinario, sono concetti distinti - e sono tali anche alla luce delle nostre radici storiche: infatti, l’idea di “privacy” nasce negli U.S.A. nel 1890, espressa come il “diritto di essere lasciati soli” (*“right to be let alone”*), strettamente legata alla protezione della propria sfera personale da intrusioni esterne. La protezione dei dati, pur essendo un sottoinsieme del concetto di privacy, si riferisce alla sicurezza e all'integrità delle informazioni personali raccolte, includendo misure tecniche e organizzative per prevenire accessi non autorizzati, perdite, alterazioni o divulgazioni illecite dei dati.

La cristallizzazione dell’idea di *privacy* fu effettuata da due giuristi statunitensi, Samuel Warren e Louis Brandeis, a seguito della pubblicazione sulla Harvard Law Review dell’approfondimento “The Right to Privacy”.



HARVARD

LAW REVIEW.

VOL. IV. DECEMBER 15, 1890. NO. 5.

THE RIGHT TO PRIVACY.

"It could be done only on principles of private justice, moral fitness, and public convenience, which, when applied to a new subject, make common law without a precedent; much more when received and approved by usage."

WILLES, J., in Millar v. Taylor, 4 Burr. 2303, 2312.

THAT the individual shall have full protection in person and in property is a principle as old as the common law; but it has been found necessary from time to time to define anew the exact nature and extent of such protection. Political, social, and economic changes entail the recognition of new rights, and the common law, in its eternal youth, grows to meet the demands of society. Thus, in very early times, the law gave a remedy only for physical interference with life and property, for trespasses *vi et armis*. Then the "right to life" served only to protect the subject from battery in its various forms; liberty meant freedom from actual restraint; and the right to property secured to the individual his lands and his cattle. Later, there came a recognition of man's spiritual nature, of his feelings and his intellect. Gradually the scope of these legal rights broadened; and now the right to life has come to mean the right to enjoy life,—the right to be let alone; the right to liberty secures the exercise of extensive civil privileges; and the term "property" has grown to comprise every form of possession—intangible, as well as tangible.

Thus, with the recognition of the legal value of sensations, the protection against actual bodily injury was extended to prohibit mere attempts to do such injury; that is, the putting another in

L’interesse degli autori scaturì a seguito di alcune pubblicazioni effettuate dall’*Evening Gazette* di Boston contenenti immagini ritraenti feste private della borghesia e dettagli sulla vita privata (e coniugale) di diversi individui, tra i quali la Signora Mabel Bayard, moglie dello stesso Warren e figlia di Thomas F. Bayard, senatore del Delaware.

Warren e Brandeis si trovarono a ragionare su quali informazioni concernenti la vita di un individuo fossero catalogabili come “di pubblico dominio”, e quali meritassero invece una tutela dall’ingerenza di terzi.

L’articolo fu di enorme importanza storica, coniugando gli insindacabili principi costituzionali della libertà di espressione e di stampa con l’effettiva rilevanza della notizia: il diritto alla privacy, pertanto, può essere esercitato salvo nei casi in cui la pubblicazione di informazioni personali non sia supportata da “pubblico o generale interesse”[1].

Il salto logico effettuato dai giuristi fu quello di andare oltre gli istituti di *libel* e *slander* (rispettivamente diffamazione mediante pubblicazioni scritte e diffamazione mediante mezzi orali), che per configurarsi richiedevano la realizzazione di un pregiudizio per l’individuo nelle proprie relazioni con le altre persone, concernendo quindi il solo danno alla reputazione[2].

---

[1] “The right to privacy does not prohibit any publication of matter which is of public or general interest.” (Citazione tradotta da lingua inglese) Samuel D. Warren, Louis D. Brandeis, “The Right To Privacy”, pag. 214 - Boston, December 1890

[2] “It deals only with damage to reputation, with the injury done to the individual in his external relations to the community, by lowering him in the estimation of his fellows.” (Citazione tradotta da lingua inglese) Samuel D. Warren, Louis D. Brandeis, “The Right To Privacy”, pag. 214 - Boston, December 1890

Furono necessari 19 anni per vedere emergere un concetto simile nel contesto europeo.
Nel 1909, Étienne-Ernest-Hippolyte Perreau scrisse *Les Droits de la Personnalité*, definendo diritti patrimoniali e diritti della personalità: questi ultimi furono classificati in 3 categorie:

- Rispetto dell'individuo stesso
- Rispetto dell'individuo come membro della famiglia
- Rispetto dell'individuo come cittadino dello Stato

Si suggerisce di rivolgere particolare attenzione al primo punto, ulteriormente suddiviso in 3 "gruppi":

- Diritto ad essere riconosciuti come un "individuo distinto dagli altri", mediante, ad esempio, il diritto al nome
- Diritto all'integrità personale dal punto di vista fisico, di salute, oltre al diritto alla vita.
- Diritto alla protezione della propria "individualità morale", concernente onore, libertà e proprietà intellettuale. Perreau menziona, senza tuttavia effettuare approfondimenti, il concetto del diritto ad "organizzare la propria vita privata".

Seppur non immediatamente accettata in Francia, la presente distinzione ha indubbiamente posto le basi ad ulteriori studi e sviluppi in merito alla delicata tematica della *privacy*[3].

Vero è che in Europa la visione predominante del tema si concentra principalmente sul concetto di protezione dei dati, originariamente derivante dal timore di una profilazione potenzialmente discriminatoria: negli anni della Seconda Guerra Mondiale, il governo dei Paesi Bassi istituì un registro anagrafico contenente dati identificativi dei cittadini quali nomi, residenza, informazioni concernenti identità economica, sociale e culturale.
A seguito dell'occupazione nazionalsocialista del territorio, tale registro cadde in mano nemica e fu utilizzato per facilitare l'identificazione di milioni di cittadini a fini persecutori sulla base delle proprie origini etniche e convinzioni religiose[4].

Alla luce di ciò, fu fondamentale garantire una protezione agli individui da ingerenze da parte di autorità pubbliche nell'ambito della propria vita privata e delle informazioni personali, con un chiaro inciso dell'articolo 8 della Convenzione Europea dei Diritti dell'Uomo (Consiglio d'Europa, 1950), il quale recita: *"Ogni persona ha diritto al rispetto della propria vita privata e familiare, del proprio domicilio e della propria corrispondenza."*. Sostiene poi, al secondo comma: *"Non può esservi ingerenza di una autorità pubblica nell'esercizio di tale diritto a meno che tale ingerenza sia prevista dalla legge e costituisca una misura che, in una società democratica, è necessaria alla sicurezza nazionale, [...] o alla protezione dei diritti e delle libertà altrui."*.

Le prime "linee" raffiguranti il concetto di privacy in Italia furono tracciate negli anni '50: prima con una sentenza del Tribunale di Roma, che sostenne come il diritto alla riservatezza si esplicasse nel *"divieto di qualsiasi ingerenza estranea nella sfera della vita privata della persona, e di qualsiasi indiscrezione da parte di terzi, su quei fatti o comportamenti personali che, non pubblici per loro natura, non sono destinati alla pubblicità delle persone che essi riguardano"*[5].

---

[3] The Development of the Theory of the Right to Privacy in France, Wencelas J. Wagner - Indiana University School of Law, 1971
[4] NIOD Institute for War, Holocaust and Genocide Studies, www.niod.nl
[5] Tribunale di Roma, 14 settembre 1953

Fu necessario attendere fino all'anno 1996, con la legge 675 (31 dicembre 1996 che recepisce in larga misura la direttiva 95/46/CE), al fine di poter offrire un'adeguata tutela alle informazioni personali delle persone fisiche. Le disposizioni prevedono l'adozione di una serie di adempimenti finalizzati a garantire i principi di correttezza, liceità e trasparenza nel trattamento dei dati personali, imponendo obblighi di informazione all'interessato in merito al trattamento di dati che lo riguardino (indicando, per esempio: finalità, natura dei dati, ambito di comunicazione e diffusione dei dati, eventuali trasferimenti verso Paesi non appartenenti all'Unione Europea, presenza di banche dati cui si riferisce il trattamento, etc.). Importante è sottolineare come tale obbligo di notificazione sussista solo qualora *"il trattamento, in ragione delle relative modalità o della natura dei dati personali, sia suscettibile di recare pregiudizio ai diritti e alle libertà dell'interessato [...]"*[6].

Il decisivo passaggio dal concetto di privacy a quello di protezione dei dati personali nel presente elaborato viene indubbiamente qui rappresentato, con particolare peso al concetto di anonimizzazione: la norma all'art. 13 (comma 1, punto c, 2) sancisce, infatti, come l'interessato abbia il diritto di ottenere [...] la cancellazione, **la trasformazione in forma anonima** o il blocco dei dati trattati in violazione di legge [...].

Non si può tuttavia non sottolineare come il contesto storico dell'emanazione dei presenti atti rappresenti solamente gli albori della massiva diffusione che internet e le nuove tecnologie hanno avuto negli anni a seguire; oltre a ciò, le divergenze nelle legislazioni nazionali, scaturite dall'utilizzo dello strumento della *direttiva*, hanno enfatizzato la necessità di uniformare il diritto comunitario con un atto che potesse fungere da *Magna Charta* in ambito protezione dei dati.

L'apice dell'elaborazione normativa fu quindi segnato dal più rilevante atto comunitario sul tema: il Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016, denominato Regolamento Generale sulla Protezione dei Dati (RGPD come acronimo in lingua italiana, GDPR in lingua inglese, più comunemente utilizzato). Il GDPR nasce dalla necessità di armonizzazione e certezza giuridica nella *data protection*, coniugando sviluppi tecnologici ed economici alle esigenze di tutela, sempre più rilevanti, per le persone fisiche.

Si comprenderà, tuttavia, come il regolamento non abbia voce in capitolo in merito ai dati anonimizzati: è comunque da ritenersi importante tracciare un quadro generale dei principi in esso contenuti, sottolineando come, qualora l'anonimizzazione non venisse effettuata secondo adeguate misure tecniche ed organizzative, i dati coinvolti possano rientrare nell'ambito di applicazione dell'atto.

[6] Art. 7. Notificazione - Legge n. 675 del 31 dicembre 1996. www.garanteprivacy.it

Il “dato personale” nel GDPR

Approfondire la tematica concernente la protezione dei dati, con particolare attenzione alle misure tecniche ed organizzative a garanzia di questa, senza quantomeno introdurre una definizione di “dato personale” (e, chiaramente, di “dato sensibile” o “particolare”), risulterebbe priva di importanti concetti e, senza dubbio, incompleta.

Alla luce dell’art. 4 del regolamento, per potersi definire “personale”, un dato deve essere un’informazione concernente una persona identificata o identificabile in modo diretto (ad es. nome e cognome) o indiretto (ad es. codice fiscale, indirizzo IP, numero di targa).

Rientrano nell’ambito di applicazione del GDPR anche quelle informazioni che descrivano l'individuo in maniera tale da consentirne l'identificazione tramite l’acquisizione di ulteriori dati (v. *pseudonimizzazione*).

Speciale preoccupazione destano i cosiddetti dati “particolari”, sottoinsieme della categoria dei dati personali, e sono definiti dall’articolo 9 del regolamento come *“dati personali che rivelino l'origine razziale o etnica, le opinioni politiche, le convinzioni religiose o filosofiche, o l'appartenenza sindacale, nonché [...] dati genetici, dati biometrici intesi a identificare in modo univoco una persona fisica, dati relativi alla salute o alla vita sessuale o all'orientamento sessuale della persona.”*.

Al giorno d’oggi, è quasi impossibile rendersi conto dell’ingente quantità di dati che noi stessi condividiamo, consapevolmente e non: dalle più delicate tematiche come una visita medica, ai più semplici atti della vita quotidiana come un acquisto in negozio o un prelievo ad un *ATM*, o la condivisione di informazioni personali sui social media. Questo flusso continuo di dati è diventato una parte essenziale della nostra interazione con il mondo digitale e concerne entrambe le tipologie sopra descritte.

Ogni volta che interagiamo con la tecnologia, lasciamo una traccia, definita *digital footprint*: le nostre preferenze, i nostri acquisti, i nostri spostamenti: a tal proposito, l’implementazione di tecniche che garantiscano una protezione adeguata dei dati è imprescindibile; tra queste, vi sono *pseudonimizzazione* ed *anonimizzazione*.

Capitolo 2: Anonimizzazione e Pseudonimizzazione nel GDPR

1.Definizioni e differenze

Nel presente capitolo verranno illustrati i principi cardine di due importanti misure mirate a minimizzare il rischio di identificazione di una persona fisica (o, nel caso dell'anonimizzazione, azzerarlo).

Sarà possibile comprendere in che modo i dati sottoposti a tali procedure subiscano variazioni in merito all'ambito applicativo del GDPR, delineandone vantaggi e svantaggi.

Fondamentale fonte per la presente tematica è il regolamento stesso, al Considerando 26, il quale recita: *"I dati personali sottoposti a pseudonimizzazione [...] dovrebbero essere considerati informazioni su una persona fisica identificabile. [...] Il presente regolamento non si applica pertanto al trattamento di tali informazioni anonime, anche per finalità statistiche o di ricerca."*.

Al fine di comprendere al meglio la suddetta distinzione, definiamo i due concetti.

Anonimizzazione e pseudonimizzazione

L'*anonimizzazione* è il processo mediante il quale l'identità di un individuo non può essere più determinata, né direttamente né indirettamente, dai dati che ne sono stati soggetto.

In altre parole, se un dato non può più essere associato ad una persona fisica, non è considerato "personale" e, di conseguenza, non rientra nel campo di applicazione del regolamento.

Si tratta di un processo **irreversibile**. Una volta che i dati vengono anonimizzati, non è più possibile risalire all'identità dell'individuo, né attraverso i dati stessi, né mediante l'uso di ulteriori informazioni. L'anonimizzazione offre indubbiamente un livello di sicurezza molto elevato, ma implica degli svantaggi: i dati anonimizzati perdono parte del loro "valore", in quanto non possono più essere utilizzati per analisi che richiedono l'identificazione dei soggetti (si pensi a finalità commerciali o mediche).

L'anonimizzazione è, tuttavia, particolarmente utile in contesti dove dati aggregati risultino sufficienti per gli scopi analitici, e nei casi in cui non sia necessario mantenere alcuna capacità di risalire all'identità dei singoli individui: ciò li rende ugualmente strumentali al raggiungimento di determinate finalità, senza però dover adempiere agli importanti vincoli imposti dal GDPR.

Con la *pseudonimizzazione*, d'altra parte, i dati vengono resi non più attribuibili ad un interessato **se non con l'utilizzo di informazioni aggiuntive**, non facendo quindi venir meno il presupposto di identificabilità della persona fisica e rendendo tale processo **reversibile**.

Il regolamento, all'articolo 4(5), richiede che le suddette informazioni aggiuntive siano *"conservate separatamente e soggette a misure tecniche e organizzative intese a garantire che tali dati personali non siano attribuiti a una persona fisica identificata o identificabile"*.

Al fine di comprendere al meglio tale distinzione, è utile presentare un esempio pratico.
Si supponga che un ospedale stia conducendo uno studio su un particolare trattamento innovativo per una determinata patologia.

I ricercatori hanno chiaramente necessità di accedere ai dati dei pazienti, come risultati di test, informazioni sulle terapie e le risposte al trattamento stesso, proteggendone l'identità.

Si supponga inoltre che i dati raccolti dai pazienti che hanno deciso di sottoporsi al trattamento riguardino nome, data di nascita, genere, indirizzo di residenza, codice fiscale, cartelle cliniche contenenti informazioni su condizioni pregresse e familiarità mediche.

Al fine di garantire un adeguato livello di sicurezza ai dati degli individui, a ciascun paziente viene assegnato un codice alfanumerico univoco casualmente generato, in sostituzione alle generalità della persona o dei propri dati medici.
Ogni informazione viene conservata su database interni alla struttura e, in ottemperanza al regolamento, i dati personali sopra citati vengono archiviati in basi di dati separate e sicure, sfruttando chiavi di accesso o strumenti come la *crittografia*. Le stesse chiavi vengono protette da accessi non autorizzati, rendendole accessibili unicamente dal personale autorizzato.

A tal punto, è possibile chiedersi: è stata effettuata una corretta pseudonimizzazione dei dati dei pazienti? Le misure di sicurezza sono comunque in grado di garantire ai ricercatori un adeguato grado di utilizzabilità dei dati stessi, alla luce delle finalità esposte?
In effetti, la risposta ad entrambi i quesiti appare affermativa: i ricercatori possono utilizzare i dati pseudonimizzati per analizzare l'efficacia del trattamento, senza svelare l'identità dei pazienti. I dati comprendono tutti gli aspetti rilevanti per lo studio ma senza alcuna informazione direttamente identificativa.
L'adeguata protezione viene garantita dal fatto che, anche nel caso in cui i dati pseudonimizzati venissero compromessi o resi accessibili a terzi non autorizzati, l'identità dei pazienti non può essere facilmente ricostruita senza accesso alle informazioni aggiuntive (quali chiavi di decodifica);

Com'è facilmente intuibile, il processo descritto non risulta essere irreversibile: infatti, qualora fosse necessario risalire all'identità di un paziente (ad esempio, per un follow up medico), il personale autorizzato può accedere alle chiavi di decodifica e ricollegare lo pseudonimo (il codice univocamente generato) all'identità reale dell'individuo.

Si supponga ora che, nella medesima fattispecie, si opti per la completa **anonimizzazione** dei dati dei pazienti: le informazioni identificative vengono rimosse completamente, eliminando nomi, indirizzi, documenti medici e qualsiasi altro dato che potrebbe essere utilizzato per identificare gli individui.
Al fine di ridurre ulteriormente il rischio di identificazione, i dati vengono generalizzati e aggregati: ad esempio, invece di conservare la data di nascita di un paziente, si opta per l'utilizzo di una fascia d'età.

I ricercatori possono utilizzare i dati anonimizzati per analizzare l'efficacia dei trattamenti, ma non vi è alcuna informazione che permetta di risalire all'identità dei pazienti: ciò comporta vantaggi e svantaggi.

I vantaggi sono, com'è intuibile, la possibilità di garantire il massimo livello di protezione della privacy; peraltro, essendo dati anonimi, non sono soggetti alle restrizioni e ai requisiti del GDPR.
Ulteriore importante vantaggio è rappresentato dall'impossibilità di identificazione dei pazienti anche nel caso di *breach* (accesso non autorizzato ai dati), problema che si pone nel caso di pseudonimizzazione (nel caso in cui un terzo non autorizzato ottenga accesso alle chiavi di decodifica, può risalire all'identità di qualsiasi paziente).

Una totale anonimizzazione però, in questo caso, presenta importanti svantaggi:
- Non è più possibile contattare i pazienti per ulteriori controlli periodici;

- La generalizzazione e l'aggregazione dei dati possono ridurre il dettaglio delle analisi (introdurre fasce d'età può influire negativamente sulla precisione dei rilievi)
- Qualora emergessero ulteriori quesiti in futuro sull'efficacia del trattamento, non sarà possibile ottenere informazioni dettagliate sui singoli pazienti.

In conclusione, ad un alto livello di protezione consegue il costo della perdita di tracciabilità e della limitazione delle capacità analitiche, rappresentando pertanto un *trade-off* tra tra massima sicurezza e protezione della privacy da un lato, e la flessibilità e l'utilità dei dati per analisi approfondite e future esigenze di ricerca dall'altro.

Capitolo 3: Tecniche di Anonimizzazione dei Dati

Una corretta anonimizzazione dei dati non può prescindere dall'applicazione di tecniche in grado di garantire la protezione della privacy degli individui, mantenendo il più possibile il valore e l'utilità delle informazioni. Le tecniche di anonimizzazione, infatti, devono essere in grado di bilanciare in maniera efficace la necessità di nascondere l'identità delle persone con la conservazione delle proprietà statistiche e analitiche.

Nel presente capitolo si analizzeranno alcune tecniche diffuse, finalizzate a minimizzare il rischio di re-identificazione dell'individuo.

Prima di ciò, tuttavia, è importante sfatare un mito: è un pensiero diffuso che una tecnica come la crittografia rientri tra le metodologie utili ad anonimizzare le informazioni; tuttavia, questa non soddisfa assolutamente il principale criterio che garantisce la completa anonimità di un dato, ovverosia l'irreversibilità delle procedure.
La crittografia utilizza algoritmi matematici e logici per trasformare i dati leggibili (definiti *plaintext*) in una forma cifrata (*ciphertext*, o testo cifrato) mediante chiave di cifratura, garantendo che solo i possessori di suddetta chiave siano in grado di applicarla al testo cifrato per ri-ottenere il dato originario. Ciò garantisce indubbiamente un ottimo livello di sicurezza nel trattamento e nel controllo degli accessi ai dati ed è infatti un metodo utilizzato ampiamente in quasi tutte le comunicazioni odierne, ma com'è intuibile, non si tratta di un processo irreversibile. Attacchi sofisticati possono essere in grado di ottenere le chiavi di *decrypt* o sfruttare vulnerabilità nei sistemi per accedere ai dati cifrati, facendo pertanto sussistere un serio rischio di re-identificazione.
Può quindi la crittografia definirsi pseudonimizzazione? Durante la stesura della presente tesi, ho avuto modo di riscontrare opinioni contrastanti: chi sostiene che la risposta sia affermativa, in quanto tecnica reversibile e che richiede ulteriori informazioni (chiave di decifratura) per poter risalire all'identità dell'individuo, e chi le ritiene due tecniche che, pur mirando al medesimo obiettivo di privacy, differiscono dal punto di vista sostanziale. Mi ritrovo ad appoggiare maggiormente quest'ultimo orientamento, in quanto l'importante e non trascurabile differenza tra le tecniche è rappresentata dall'accessibilità dei dati: nella pseudonimizzazione il dato stesso non viene reso inintelligibile, ma viene unicamente minimizzato il rischio di re-identificazione, mentre nella crittografia l'oscuramento del dato avviene *by default*, rendendo l'informazione non comprensibile se non mediante una chiave di accesso.
Nulla vieta di combinare la tecnica di pseudonimizzazione con la cifratura dei dati: tale integrazione può portare ad un livello ancora maggiore di sicurezza delle informazioni.

1.ISO/IEC 27559:2022 e le procedure standardizzate

Conseguentemente ad una critica lettura delle normative ad oggi vigenti in merito alla protezione dei dati personali, non risulta presente alcuna descrizione tecnica sulle misure da adottare. Tant'è vero che lo stesso GDPR, in ottemperanza al principio di *accountability*, prevede l'adozione di misure tecniche e organizzative per rendere il trattamento dei dati sicuro, lasciando apparentemente libera valutazione al titolare in merito all'effettiva applicazione delle metodologie.

Vero è che, per quanto questa potrebbe essere considerata una mancanza da parte di alcuni, un corpo normativo del genere ha la funzione di fissare i più importanti principi e obiettivi secondo un approccio risk-based, piuttosto che descrivere nei dettagli i passaggi tecnici da applicare. In effetti, essendo il regolamento (e non solo) applicabile ad un indefinito numero di casistiche, alcune metodologie possono non essere adatte al caso concreto in quanto non sufficienti o addirittura eccessive.
A tal proposito, l'utilizzo di procedure standardizzate garantisce un adeguato livello di sicurezza nelle informazioni, oltre all'ottemperanza ai criteri fissati dalle normative vigenti.
ISO/IEC 27559:2022 definisce una serie di pratiche per una corretta anonimizzazione dei dati, riferendosi al titolare del trattamento e offrendo linee guida affinché questi possa decidere in merito alla misura maggiormente opportuna visto il caso concreto, qualora si trovasse nella situazione di dover trasmettere a terzi informazioni personali di altri individui.

La norma, sin dalle prime battute, sottolinea come le misure presenti nel documento siano applicabili ad ogni tipo di organizzazione, pubbliche o private che siano, comprendendo le associazioni non-profit e tutti i titolari del trattamento i quali implementino misure di de-identificazione dei dati[1].

Interessante notare come non venga mai utilizzata l'espressione "anonimizzazione", in quanto adottata nel tempo da diversi legislatori con differenti connotazioni, sostituendola con "de-identificazione": dal punto di vista sostanziale non vi sono differenze, ma data la natura generica dello standard si è preferito utilizzare una terminologia in grado di non generare equivoci in sede applicativa.

ISO/IEC 27559:2022 sottolinea come la de-identificazione sia l'insieme di processi e tecniche atte a trasformare dati riferibili ad una persona fisica in dati che non consentano in alcun modo l'identificazione dell'interessato, allineandosi quindi a quella comunque comunemente definibile "anonimizzazione". Tali procedure sono essenziali a mitigare i rischi di identificazione degli individui, in particolar modo nei casi di futuro riutilizzo interno all'organizzazione, di condivisione esterna (il titolare o - come vi si riferisce lo standard - il custode, mette a disposizione i dati al di fuori dalla propria organizzazione ma in un ambiente sotto il suo controllo come un centro di ricerca, un'Università convenzionata) e di rilascio esterno, dove la condivisione delle informazioni avviene in un ambiente non supervisionato.

---

[1] Da ISO/IEC 27559:2022: "This document is applicable to all types and sizes of organizations, including public and private companies, government entities, and not-for-profit organizations, that are PII controllers or PII processors acting on a controller's behalf, implementing data de-identification processes for privacy enhancing purposes."

Lo standard divide le procedure di de-identificazione in quattro fasi:

- **Analisi del contesto** (Context Assessment): ogni adeguata misura di sicurezza è certamente frutto di una corretta valutazione della tipologia di dati trattati e comunicati a terzi. È importante altresì determinare il livello di trasparenza mediante il quale le informazioni verranno condivise e, soprattutto, l'"ambiente" nel quale si intende diffonderle. Al contempo, risulta fondamentale effettuare una attenta analisi a:
    - **Minacce** (Threat Modeling): le minacce possono essere catalogate in 3 categorie:
        - **Intenzionali** (Deliberate), in cui il riferimento è ad un attacco volontario da parte di un insider dell'infrastruttura a cui il custode sta trasferendo i dati (e.g. un dipendente)
        - **Accidentali** (Accidental), dove l'identificazione di un individuo all'interno di un dataset avviene involontariamente (il destinatario riconosce uno o più individui all'interno del dataset; ciò implica, per il custode, la necessità di includere una valutazione in merito alla presunta conoscenza che il destinatario dei dati abbia su di essi)
        - **Ambientali** (Environmental): si tratta di perdita e/o furto di dati personali causato da inadeguatezza o malfunzionamento dei sistemi di sicurezza informatici.
    - **Valutazione d'impatto e trasparenza** (*Transparency and Impact Assessment*): a completamento della valutazione del contesto, lo standard suggerisce di coinvolgere vari *stakeholder* del processo di trasmissione dei dati, come gli interessati stessi o le Autorità competenti in materia di privacy. È utile effettuare una valutazione d'impatto identificando, oltre ai rischi stessi, le metodologie per poterli mitigare, tramite un approccio di *privacy by design* (come menzionato nello stesso GDPR)
- **Data Assessment** (Valutazione sui Dati): lo scopo del *data assessment* è quello di analizzare e comprendere a che tipologie di informazioni il destinatario avrà accesso e quali di queste presentino il rischio di re-identificazione. Grazie a tale valutazione, è possibile effettuare decisioni in merito a quali dati anonimizzare e quali lasciare inalterati.
  A tal proposito, lo standard offre un'utile e strutturata modalità di categorizzazione dei dati: viene valutato se, all'interno del dataset, i dati rappresentino una singola unità (persona o famiglia) o un'aggregazione di più unità (e.g. popolazione di una città). Particolare attenzione deve essere prestata all'eventuale presenza di individui fragili o potenzialmente vulnerabili.
  Ciascun dato presenta degli attributi, i quali si esplicano in tutte le informazioni contenute nel dataset in grado di identificare l'individuo direttamente ed indirettamente, analizzando il loro livello di unicità (un attributo unico e non ripetuto all'interno di un dataset porta ad identificare univocamente il suo possessore) e di "sensibilità" dei dati (informazioni sulla salute dell'individuo, sulla propria situazione economica, familiare, sociale, etc.).

  Il *data assessment* prevede inoltre una quantificazione dei rischi, dividendoli in:
    - Prosecutor risk: l'avversario (*adversary*) è a conoscenza della presenza di un particolare individuo all'interno di un dataset
    - Journalist risk: l'avversario non ha contezza della presenza di un individuo all'interno di un dataset
    - Marketer's risk: non vi è finalità di riconoscimento di un singolo individuo nel dataset, bensì di identificazione di più individui possibili, ad esempio per scopi commerciali.

- **Identifiability Assessment and Mitigation** (Valutazione sull'identificabilità e mitigazione di questa): si tratta della probabilità che il rischio di re-identificazione di un individuo, grazie a dati contenuti in un dataset parzialmente o completamente divulgato, si realizzi.

Definita ***P(Id) = P(Id | R) x P(R)*** e letta come "*Probabilità di Identificazione* uguale alla *Probabilità di Identificazione data la realizzazione di un Rischio*, moltiplicata per la *Probabilità di realizzazione del Rischio stesso*", la formula rappresenta una valutazione numerica del rischio di re-identificazione e, sulla base del risultato prodotto, si è in grado di ottenere una stima sull'entità del rischio stesso.

Infatti, grazie a dei valori limite indicati dall'*Annex B* dello standard, pari al range tra 0.0005 e 0.1[2], è possibile concludere se il rischio sia o meno sopportabile (considerando la tipologia di attacco ed i gruppi di dati coinvolti).

Sulla base del risultato prodotto (ma anche, chiaramente, di valutazioni in merito al caso concreto), ISO/IEC 27559:2022 suggerisce di rimodulare la tipologia di dati trasmessi ai destinatari o effettuare delle elaborazioni su di essi finalizzate a ridurre ulteriormente il rischio di re-identificazione (tra queste, generalizzazione e campionamento).

La sicurezza dei dati non è certamente un concetto binario, bensì ciclico: infatti, è indubbiamente consigliabile effettuare un'ulteriore analisi sull'identificabilità, verificando che il risultato successivo alle suddette elaborazioni rientri nel range indicato nello standard.

- **De-Identification Governance** (Governance della De-identificazione): lo standard sottolinea come l'implementazione di una policy sulla condivisione dei dati possa essere integrata all'interno di pratiche più generali concernenti la protezione delle informazioni.
  La suddetta policy deve essere applicabile sia prima che dopo la trasmissione dei dati.
  - Prima: si ritiene utile identificare ruoli e responsabilità al fine di effettuare un trasferimento in ottemperanza alle normative vigenti, tenendo traccia dei dati coinvolti e interpellando gli *stakeholders* per garantire la concretizzazione del suddetto principio di trasparenza.
  - Dopo: successivamente alla condivisione delle informazioni, è prevista una regolare verifica in merito alle condizioni dell'ambiente nel quale i dati sono stati trasmessi, per garantire sicurezza e minimizzare il rischio di re-identificazione nel tempo (il progresso tecnologico può presentare minacce non state oggetto di valutazione al momento della *data disclosure*).

I principi descritti nel presente capitolo si possono applicare ad innumerevoli utilizzi, dal campo medico a quello commerciale; ogni contesto richiede un approccio specifico in base ai diversi rischi che possono emergere, così come differenti misure tecniche, illustrate nelle prossime pagine.

[2] Da https://www.private-ai.com/en/2023/05/17/iso-iec-27559-2022/, azienda canadese specializzata in privacy e machine learning, partner dell'Università di Toronto

2.Metodologie diffuse

Nella presente sezione verranno descritte le misure tecniche maggiormente diffuse per garantire un'adeguata anonimizzazione dei dati personali, identificando vantaggi e svantaggi.

Tecniche di offuscamento dei dati personali

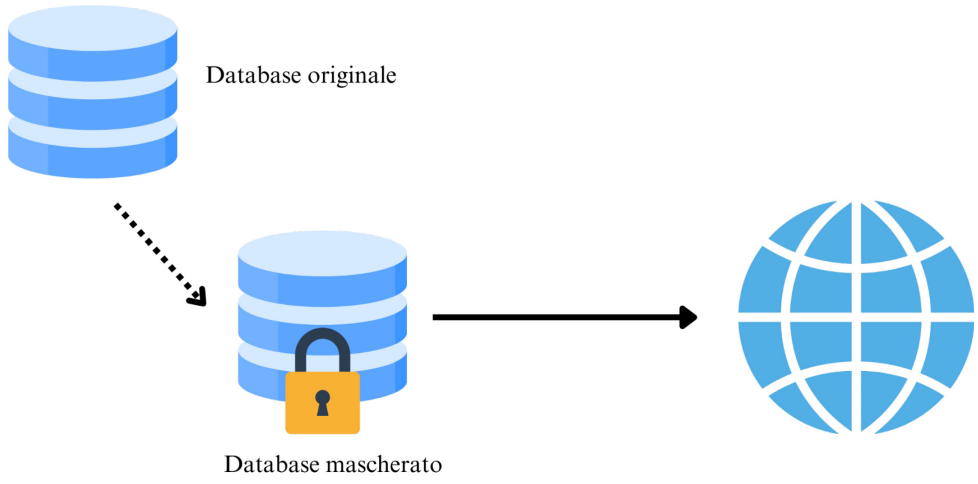
Le seguenti sono solo alcune delle tecniche più utilizzate nell'anonimizzazione dei dati personali e possono presentare variazioni a seconda del caso concreto. Com'è stato in precedenza illustrato, mantenere un adeguato livello di flessibilità in merito alle metodologie applicate garantisce un miglior adattamento al singolo contesto di applicazione.

Una tecnica particolarmente diffusa è rappresentata dal **data masking**: si tratta dell'alterazione o dell'offuscamento di alcuni valori all'interno del *dataset*, garantendone comunque accessibilità ma non permettendo di risalire al dato originario.

I dati vengono sostituiti con caratteri o informazioni artificiose, generando quindi delle informazioni che non rappresentano la realtà dei fatti, ma ugualmente in grado di preservare le caratteristiche del *dataset*.

È bene sottolineare come il *data masking* possa essere utilizzato sia nel caso di anonimizzazione che in quello di pseudonimizzazione, e :

- **Static Data Masking (SDM)**: prevede l'effettuazione di una copia dei dati nel dataset, separata dall'originale, nella quale i dati personali vengono mascherati in modo irreversibile. La copia può quindi essere trasmessa e condivisa con la consapevolezza che, a seguito di un processo di *masking* efficace svolto a monte, non sia in alcun modo possibile identificare gli individui in essa contenuti.
  La procedura può riassumersi nel seguente schema:



La copia del database a cui viene applicata la procedura di masking si può presentare così:

Nome	Reddito
Mario Rossi	€50.000
Luigi Bianchi	€20.000
Gianni Verdi	€40.000

Database originale

Nome	Reddito
MXXXX RXXXX	€50.000
LXXXX BXXXXXX	€20.000
GXXXXX VXXXX	€40.000

Database mascherato

Il dataset definito a destra contiene quindi dati anonimizzati in maniera irreversibile: nessun *attacker* può essere in grado di risalire alle informazioni personali grazie ai dati mascherati. Le semplici iniziali non rappresentano un modo per risalire ai dati personali (si immagini un dataset ben più vasto rispetto al mero esempio contenente 3 *record*. In alternativa, è possibile mascherare anche le iniziali).

L'utilità del database può non essere stata in alcun modo alterata, a seconda della finalità: nel caso in cui si stesse calcolando una media matematica del reddito percepito da una parte di popolazione, la presenza di informazioni come nome e cognome non è rilevante.

È interessante valutare se la presenza del database originale, contenente dati personali, possa in qualche modo inficiare sul carattere irreversibile del processo di anonimizzazione. In caso di accesso ai dati in chiaro l'identificazione è immediata, ma ciò non deriva da una fallacia nel processo di masking, quanto da inadeguate misure di sicurezza applicate al dataset originale, il quale dovrebbe essere isolato o, per azzerare il rischio di identificazione ed uscire dal campo di applicazione del GDPR, distrutto. Infatti, un qualsiasi malintenzionato in possesso del solo database correttamente mascherato non può in alcun modo risalire alle persone fisiche in esso contenute.

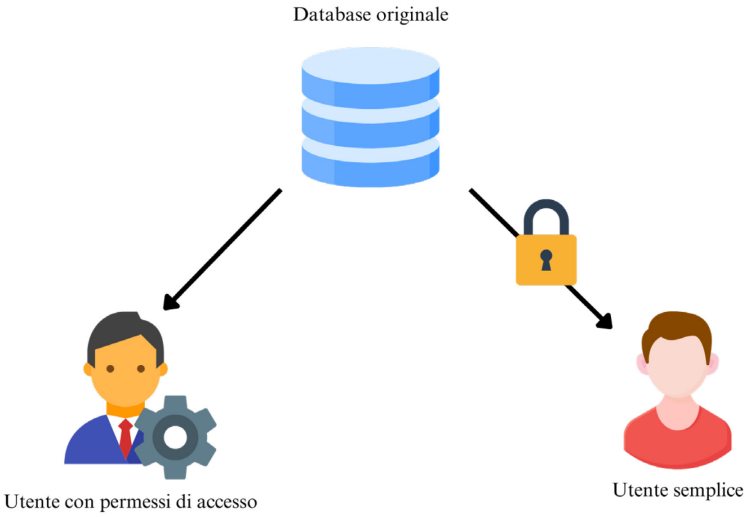
Chiari svantaggi di questa metodologia è rappresentato dalla necessità di un doppio spazio di archiviazione per la copia del database e dal tempo necessario per anonimizzare una grande quantità di dati (si pensi a dataset con centinaia di milioni di informazioni), oltre alla necessità di ri-mascheramento dei dati ad ogni modifica di questi.

- **Dynamic Data Masking (DDM)**: per sopperire agli svantaggi sopra menzionati, il mascheramento dinamico appare la soluzione più agile.
  Il mascheramento dei dati non avviene mediante copia del *dataset*, bensì in tempo reale: al momento dell'esecuzione di una *query*, il sistema di DDM intercetta i dati di risposta e applica le maschere sulla base di regole predefinite e autorizzazioni per gli accessi ai dati in chiaro. Il mascheramento, come nell'SDM, può essere completo (ogni carattere viene sostituito) o parziale (alcuni caratteri vengono sostituiti, v. tabella nell'esempio del punto precedente).

  Esempio molto comune è il mascheramento del numero della carta di credito: al momento dell'acquisto in un marketplace come Amazon.it, non di rado il metodo di pagamento viene identificato mediante le ultime 4 cifre della carta.

  In un grande marketplace, la quantità di dati personali registrata è così ingente da rendere inefficiente un mascheramento statico mediante copia dei database.

Lo schema funzionale del DDM può essere così riassunto:



Immaginiamo che entrambi gli utenti vogliano accedere al record contenente i dati di Mario Rossi, utente di sesso maschile nato l'1/1/1980 con numero di telefono 123 4567890 e reddito €50.000. Supponiamo che l'utente con permessi di accesso sia lo stesso Mario Rossi, il quale si è autenticato all'interno del sistema tramite credenziali d'accesso, e che il secondo utente non abbia permessi di lettura dei dati in chiaro ma desideri ottenere informazioni a fini statistici sul reddito medio della popolazione in base ad età e genere.
Il *record* che il database restituirà a Mario conterrà i dati in chiaro, mentre quello restituito all'utente 2 può essere:

Nome	Data di nascita	Sesso	Telefono	Reddito
XXXXX XXXXX	XX/XX/1980	M	XXX XXXXXXX	€50.000

Come si può notare, il mascheramento è stato totale sul nome e sul numero di telefono, non rilevanti a fini statistici, mentre è stato parziale sulla data di nascita per poter comunque usufruire dell'anno, e non è per niente avvenuto sugli attributi "Sesso" e "Reddito", i quali ovviamente non identificano di per sé alcun individuo.

Un tipo di mascheramento dinamico richiede quindi meno spazio di archiviazione ed un livello maggiore di flessibilità in quanto i dati vengono oscurati solo quando necessario, ma li rende anonimi solo in visualizzazione e non in archiviazione: infatti, l'elaborazione avviene sempre sulle informazioni in chiaro (garantendo peraltro accesso a queste da parte di utenti autorizzati), rientrando pertanto nel campo applicativo del GDPR.

Un secondo metodo di offuscamento dei dati personali è rappresentato dalla **generalizzazione**, la quale si può descrivere come "*il processo di selezione di un valore distintivo e di astrazione in un valore più generale e meno distintivo*[3]".
Se eseguita correttamente, si tratta di una tecnica in grado di alterare in maniera minima l'utilità del dati effettuando un'astrazione su di essi, catalogando ciascun individuo all'interno di range predefiniti nel dataset.

[3] Da https://cloud.google.com/sensitive-data-protection/docs/concepts-bucketing?hl=it

Al fine di una comprensione immediata della metodologia, si consideri il seguente esempio:
Un database contiene nomi, età, indirizzi di residenza e redditi di diversi individui. I dati sono necessari per un'indagine statistica.

Nome	Età	Residenza	Reddito
Mario Rossi	39	Via Roma 1, Milano	€50.000
Luigi Bianchi	29	Via Garibaldi 5, Napoli	€20.000
Gianni Verdi	35	Via V. Emanuele 3 , Milano	€40.000

Database originale

Età	Residenza	Reddito
31-40	Milano	€30.000 - €50.000
21-30	Napoli	€10.000 - €20.000
31-40	Milano	€30.000 - €50.000

Database generalizzato

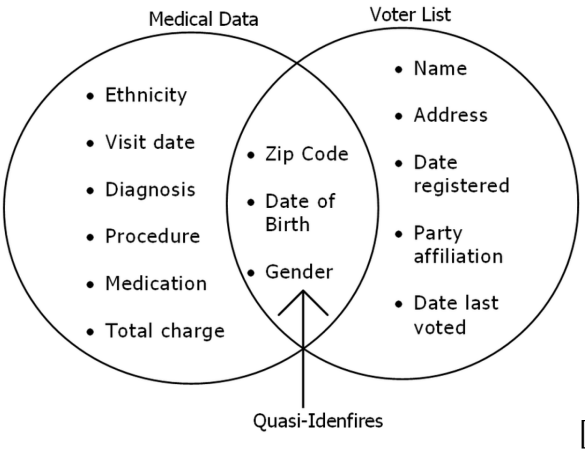
A seguito del processo di generalizzazione, l'età esatta dell'individuo ed il reddito sono stati ridotti ad un range per ridurre la precisione nei dati ed evitare possibili re-identificazioni grazie ad essi (ad esempio incrociando le informazioni con altri database) e l'indirizzo di residenza alla sola città di residenza.

Si noti l'assenza dell'attributo "Nome", in quanto diretto identificatore della persona e non necessario a fini statistici (mediante procedura denominata *soppressione*). Anche un mascheramento dell'attributo stesso avrebbe rappresentato una possibile soluzione, ma al fine di limitare lo spazio di archiviazione richiesto, si è ritenuto maggiormente adatto eliminare l'intera colonna.

Sorge spontanea, tuttavia, una domanda: com'è possibile definire il range entro il quale i dati vengono generalizzati? È chiaro come un range esageratamente ampio, per quanto preservi maggiormente la privacy degli individui, possa inficiare sull'utilità dei dati stessi (in uno studio medico per valutare l'incidenza di una patologia sulla base di fasce d'età, generalizzare tali fasce in range come 0-50 e 51+ anni può portare ad importanti imprecisioni e minare la veridicità dei risultati). È altresì vero che un range eccessivamente ristretto non sia in grado di garantire adeguati livelli di sicurezza.

A tal proposito, indispensabile è l'introduzione della **k-anonimity** (k-anonimato) e del concetto di *quasi-identifier*.

Un *quasi-identifier* è un attributo che, se combinato ad altre informazioni, può essere utilizzabile per identificare un individuo specifico.



[4]

[4] Tratto da "An improved differential privacy algorithm to protect re-identification of data", A. N K Zaman, Charlie Obimbo, Rozita A. Dara

A differenza di un identificatore diretto, come il nome completo o il codice fiscale, grazie al quale il riconoscimento di un individuo è immediato, un quasi-identificatore permette di ridurre il gruppo di individui possibili ad un gruppo più ristretto (specialmente in *dataset* di dimensioni ridotte) e addirittura identificare l'individuo univocamente tramite la conoscenza di ulteriori informazioni.

Supponiamo che un *attacker*, o *adversary*, abbia contezza del nome di una persona nell'insieme di destra dell'immagine, oltre al suo CAP di residenza e data di nascita (è quindi a conoscenza di un identificatore diretto e di due *quasi-identifiers*). L'*attacker* cerca all'interno del database contenente record medici (insieme di sinistra della stessa immagine) dati relativi a CAP e data di nascita, in quanto il campo "nome" è stato omesso dal gestore della base di dati per ragioni di privacy nel campo medico.
A seguito dell'interrogazione del database, questo ritorna un singolo record: l'*attacker* ha così ottenuto i dati della persona a cui stava mirando, avendo pertanto accesso a tutte le sue informazioni mediche.

È da sottolineare che, senza la conoscenza di informazioni personali come il nome, l'*attacker* avrebbe avuto la sola contezza dei dati medici di un individuo di un determinato genere, residente all'interno di un certo Comune e nato in una certa data, senza però conoscerne l'identità.

Si osservi peraltro che, qualora il database avesse risposto con più *record* e non una singola riga, l'identificazione univoca non sarebbe stata possibile: ciò dipende anche dalla dimensione del *dataset* (la probabilità che più dati si "assomiglino" è chiaramente maggiore in presenza di numerose persone).

Al fine valutare il livello di identificabilità degli individui in un *dataset*, introdurre il concetto di *k-anonimity* è fondamentale.

La k-anonimity fu per la prima volta descritta nel 1998 da Pierangela Samarati e Latanya Sweeney[5] ed è una proprietà altamente diffusa e sfruttata ancora oggi.

Nei dataset un errore comune è quello di rimuovere unicamente gli identificatori di un individuo (come nome e cognome, o il codice fiscale) con la convinzione che ciò rappresenti un'operazione sufficiente a preservarne la privacy. La stessa Latanya Sweeney scoprì che, mediante il solo utilizzo di data di nascita, genere e CAP (zip code), l'87% dei cittadini statunitensi presenti nei database pubblici fosse identificabile[6]. A tal proposito, ulteriori valutazioni devono essere effettuate.

Per far sì che un dataset soddisfi la k-anonimity, devono esistere almeno k individui nel dataset stesso che condividono gli stessi attributi, al fine di rendere non univoco un eventuale riconoscimento.

Per definirla tecnicamente, sia $T(A_1, \ldots, A_n)$ una tabella con una serie di attributi, e QI un quasi-identificatore o un gruppo di questi. La tabella T soddisfa la proprietà di k-anonimity rispetto a QI se ogni sequenza di valori in T[QI] (record della tabella contenenti il quasi-identificatore o gruppo di quasi-identificatori QI) appare almeno k volte in T[QI].

Si tratta di un concetto facilmente comprensibile se accompagnato da un esempio:

---

[5] Samarati Pierangela; Sweeney Latanya (1998). "Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression" (PDF). Harvard Data Privacy Lab.
[6] L. Sweeney, Simple Demographics Often Identify People Uniquely. Carnegie Mellon University, Data Privacy Working Paper 3. Pittsburgh 2000

Nome	Età	CAP	Stato
Mario Rossi	45	45100	Sposato/a
Luigi Bianchi	38	45100	Divorziato/a
Gianni Verdi	66	35100	Separato/a
Andrea Neri	70	35100	Sposato/a

Database originale

	Età	CAP	Stato
Gruppo di equivalenza 1	31-50	45100	Sposato/a
	31-50	45100	Divorziato/a
Gruppo di equivalenza 2	51-70	35100	Separato/a
	51-70	35100	Sposato/a

Database che soddisfa la k-anonimity

Il database originale contiene dati di 4 individui quali nome, età, CAP di residenza e stato civile. Al fine di rendere ogni record anonimo, i passaggi svolti sono i seguenti:

- Rimozione del campo "Nome" in quanto identificatore diretto;
- Generalizzazione dell'età in un range;

- Nessuna modifica al CAP di residenza in quanto, se si volesse risalire ad uno tra, ad esempio, Mario Rossi e Luigi Bianchi, ora non sarebbe possibile: nel database anonimizzato, infatti, appaiono due individui dell'età compresa tra 31 e 50 anni, residenti all'interno del Comune di Rovigo (45100), di cui uno sposato e l'altro divorziato, senza poter individuare chi sia chi (ed il medesimo discorso si applica al secondo gruppo di equivalenza).

(Un gruppo di equivalenza è un insieme di record che condividono i medesimi quasi-identificatori, in questo caso età e CAP).

Il database di cui sopra è quindi definito come *2-anonymous*, in quanto ogni combinazione dei quasi-identificatori appare almeno due volte (in altre parole, ci sono almeno 2 persone per gruppo di equivalenza).

Es.2:

Nome	Età	CAP	Stato
Mario Rossi	45	45100	Sposato/a
Luigi Bianchi	38	45100	Divorziato/a
Gianni Verdi	66	35100	**Separato/a**
Andrea Neri	70	35100	**Separato/a**

Database originale

	Età	CAP	Stato
Gruppo di equivalenza 1	31-50	45100	Sposato/a
	31-50	45100	Divorziato/a
Gruppo di equivalenza 2	51-70	35100	**Separato/a**
	51-70	35100	**Separato/a**

Database che **NON** soddisfa la k-anonimity

Supponiamo ora, nel medesimo dataset, che Andrea Neri si sia separato. Se applicassimo gli stessi passaggi per anonimizzare i dati e prevenire il rischio di re-identificazione, noteremmo una mancanza di diversità negli attributi sensibili del gruppo di equivalenza 2: infatti, pur non essendo a conoscenza di quale dei due record si riferisca ad Andrea Neri, sapremmo che il suo stato civile è quello di "Separato".
Per sopperire alla presente problematica, può essere necessario rimuovere sia l'età che il CAP, oppure generalizzare ulteriormente accorpando i dati del primo gruppo a quelli del secondo (rischiando in entrambi i casi di perdere parte l'utilità delle informazioni qualora stessimo, per mero esempio, svolgendo uno studio sul numero di individui separati o divorziati nelle città di Rovigo e Padova).

È importante sottolineare che gli esempi qui proposti ritraggono un numero limitato di individui per

ovvie questioni di praticità. Applicare i presenti concetti a dataset con numerosità più elevate può quindi risultare più efficace (è chiaro come non esistano solamente due individui tra i 51 e i 70 anni nel Comune di Padova, pertanto la generalizzazione è meglio applicabile alla realtà dei fatti); tuttavia, è altresì vero che sia possibile risalire a dati sensibili tramite *omogeneity attack*, attacco che sfrutta esattamente la non diversità delle informazioni personali qui illustrata (nonché principale debolezza della proprietà).

Estensione della k-anonimity è la **l-diversity**, proprietà che sostiene come sia necessario disporre di almeno un numero "l" di valori distinti nei campi contenenti attributi sensibili all'interno di ciascuna classe o gruppo di equivalenza.

Se volessimo ottenere un database che soddisfi la proprietà, dovremmo riutilizzare la prima versione del dataset contenente informazioni sullo stato civile di 4 individui:

Nome	Età	CAP	Stato
Mario Rossi	45	45100	Sposato/a
Luigi Bianchi	38	45100	Divorziato/a
Gianni Verdi	66	35100	Separato/a
Andrea Neri	70	35100	Sposato/a

Database originale

	Età	CAP	Stato
Gruppo di equivalenza 1	31-50	45100	Sposato/a
	31-50	45100	Divorziato/a
Gruppo di equivalenza 2	51-70	35100	Separato/a
	51-70	35100	Sposato/a

Database che soddisfa la k-anonimity **E** la l-diversity

Com'è stato illustrato in precedenza, il dataset è 2-anonimo in quanto caratterizzato da classi di equivalenza contenenti 2 record e, a ben vedere, sia la prima che la seconda classe di equivalenza sono *2-diverse* (eng.) in quanto ciascuna classe contiene 2 valori distinti nell'attributo sensibile "Stato".

La l-diversity, per quanto chiara evoluzione della k-anonimity, presenta una criticità tutt'altro che irrilevante: la mancanza di un'analisi semantica degli attributi.

A tal proposito, un'ulteriore miglioria viene presentata dal concetto di **t-closeness**.

Si supponga ora che Andrea Neri subisca un ulteriore cambio di stato civile, passando allo stato di "Divorziato/a".

Nome	Età	CAP	Stato
Mario Rossi	45	45100	Sposato/a
Luigi Bianchi	38	45100	Divorziato/a
Gianni Verdi	66	35100	Separato/a
Andrea Neri	70	35100	Divorziato/a

Database originale

	Età	CAP	Stato
Gruppo di equivalenza 1	31-50	45100	Sposato/a
	31-50	45100	Divorziato/a
Gruppo di equivalenza 2	51-70	35100	Separato/a
	51-70	35100	Divorziato/a

Database che soddisfa la k-anonimity **E** la l-diversity, ma **NON** la t-closeness

Analizziamo il dataset per gradi, riassumendo brevemente le proprietà finora illustrate:

1. Soppressione dell'attributo "Nome" in quanto diretto identificatore della persona;
2. Generalizzazione dell'attributo "Età" per minore precisione in caso di tentativo di re-identificazione;

3. Nessuna modifica in merito ai CAP.

Proprietà:

1. k-anonimity: soddisfatta, in quanto ciascun gruppo di equivalenza presenta 2 record;
2. l-diversity: soddisfatta, in quanto sono presenti almeno due attributi sensibili differenti in ciascuna classe di equivalenza;
3. t-closeness: **non** soddisfatta nel gruppo di equivalenza 2. Perché?

Analizzando la semantica dell'attributo "Stato" nel gruppo n.2, è possibile notare che in entrambi i casi ci si riferisce ad una situazione coniugale negativa.

La *t-closeness* è una proprietà in grado di ridurre la granularità (precisione, livello di dettaglio) dei dati, al contempo riducendo i rischi di identificazione.
Principio introdotto per la prima volta da Ninghui Li, Tiancheng Li e Suresh Venkatasubramanian[7], si sostiene che una classe di equivalenza soddisfi la proprietà t-closeness se la distanza tra la distribuzione di un attributo sensibile della classe stessa e la distribuzione dell'attributo nell'intero dataset non differisca per più di un certo parametro limite *t*[8].

Tale distanza può essere calcolata sia nel caso di attributi numerici che non numerici (si pensi ad un dato come lo stipendio di un individuo nel primo caso e una patologia nel secondo) tramite il metodo denominato *Earth Mover's Distance (EMD)*.

Data la differenza di applicazione a seconda della natura dell'attributo (numerico e non), risulta utile introdurre un secondo dataset, con più elevata numerosità, tratto e tradotto dal documento di Ninghui Li, Tiancheng Li, e Suresh Venkatasubramanian:

Nome	Età	CAP	Reddito	Patologia
Alice	29	47677	3.000	Ulcera gastrica
Bob	22	47602	4.000	Gastrite
Charly	27	47678	5.000	Carcinoma gastrico
Dave	43	47905	6.000	Gastrite
Eve	52	47909	11.000	Influenza
Ferris	47	47906	8.000	Bronchite
George	30	47605	7.000	Bronchite
Harvey	36	47673	9.000	Polmonite
Iris	32	47607	10.000	Carcinoma gastrico

Dataset originale

Anonimizziamo il dataset applicando le proprietà descritte.

1. Rimozione dell'attributo nome in quanto identificatore diretto;
2. Introduzione di range di età;
3. Mascheramento parziale del CAP sfruttando le cifre comuni nei vari attributi al fine di riorganizzare i record in gruppi di equivalenza.

[7][8] Li, Ninghui; Li, Tiancheng; Venkatasubramanian, Suresh (2007). "T-Closeness: Privacy Beyond k-Anonymity and l-Diversity". t-Closeness: Privacy beyond k-anonymity and l-diversity

	Età	CAP	Reddito	Patologia
Cl. eq. 1	20-29	476**	3.000	Ulcera gastrica
	20-29	476**	4.000	Gastrite
	20-29	476**	5.000	Carcinoma gastrico
Cl. eq. 2	40-59	4790*	6.000	Gastrite
	40-59	4790*	11.000	Influenza
	40-59	4790*	8.000	Bronchite
Cl. eq. 3	30-39	476**	7.000	Bronchite
	30-39	476**	9.000	Polmonite
	30-39	476**	10.000	Carcinoma gastrico

Dataset che soddisfa k-anonimity ed l-diversity

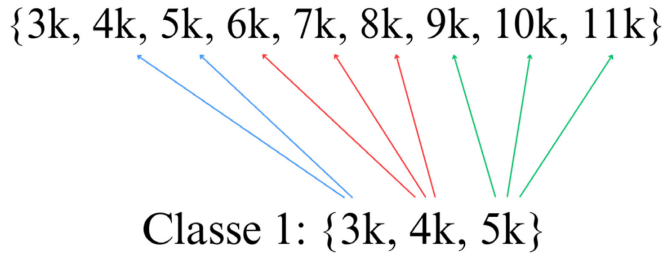
Come menzionato in precedenza, la proprietà l-diversity non tiene in considerazione l'aspetto semantico degli attributi: nel presente dataset ed in particolar modo nella prima classe di equivalenza, grazie ai *quasi-identifiers* "Età" e "CAP", è possibile sapere che tutti 3 gli individui (Alice, Bob e Charly) hanno un reddito relativamente basso e soffrono di problematiche all'apparato digerente.

Al fine di minimizzare ulteriormente il rischio di re-identificazione, applichiamo i metodi indicati dal principio di t-closeness iniziando dall'attributo numerico "Reddito".

1. Trascriviamo la distribuzione ordinata del reddito dell'intero database (utilizzo "k" per indicare le migliaia per una scrittura più compatta)
   {3k, 4k, 5k, 6k, 7k, 8k, 9k, 10k, 11k}

2. Consideriamo a titolo esemplificativo le prime due classi di equivalenza, le quali comprendono rispettivamente i valori **3k, 4k, 5k** e **6k, 11k, 8k**.

3. Applichiamo la seguente formula, nonché semplificazione della ben più complessa EMD:
   $$Do = \frac{|i - j|}{n - 1}$$
   dove "Do" rappresenta la distanza ordinata, "i" e "j" due oggetti in due diverse distribuzioni ed "n" il numero totale di oggetti nella distribuzione originale.



Confrontiamo gli attributi "Reddito" della prima classe di equivalenza con la distribuzione originale. A fini esemplificativi, il valore massimo della Classe 1 viene confrontato ai 3 valori massimi della distribuzione originale, ed il medesimo principio viene applicato al valore medio (4k) ed al valore minimo (3k), fatta salva la comparazione con se stesso in quanto

irrilevante ai fini del calcolo. Ciò permette calcoli più rapidi ed una complessità inferiore.

Calcoliamo ora la distanza assoluta tra elementi. In assenza di numeri negativi, il valore assoluto è omesso.

4-3 + 5-3 + 6-4 + 7-4 + 8-4 + 9-5 + 10-5 + 11-5 = 27 →otteniamo | i - j |

Dividiamo il risultato ottenuto per n-1 elementi della distribuzione originale, ossia 9-1 = 8.

27/8 = 3.375

L'apparizione di ciascun elemento all'interno della distribuzione originale è equiprobabile ed è pari ad 1/9, per cui:

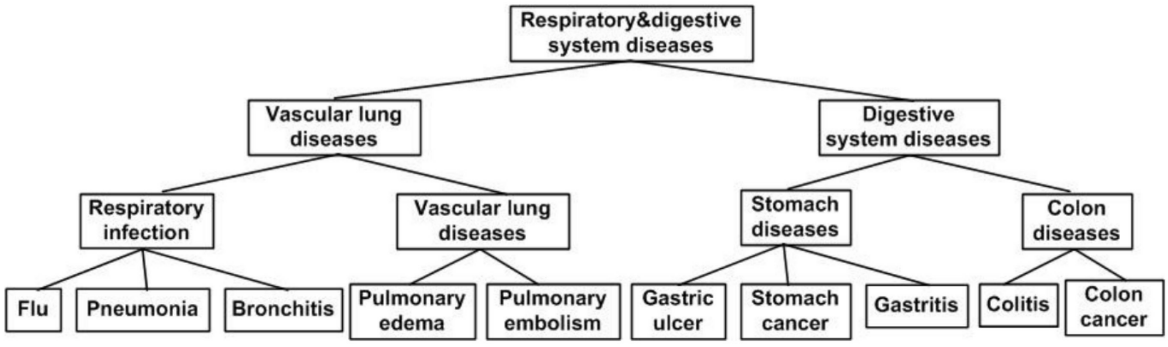
3.375/9 = 0.375

Tale valore è chiamato *optimal mass flow*, ossia una misura dello sforzo medio necessario per trasformare una distribuzione in un'altra, in base alla distanza tra i valori. Un valore basso indica che lo sforzo è minore, in quanto gli elementi risultano simili tra di loro.

Applicando il medesimo procedimento anche alla classe di equivalenza 2, è possibile ricavare il valore **0,167**[9]. La distribuzione della presente classe è, quindi, più vicina all'originale.

Per quanto concerne le distribuzioni di dati non numerici, la valutazione avviene tramite una struttura ad albero.

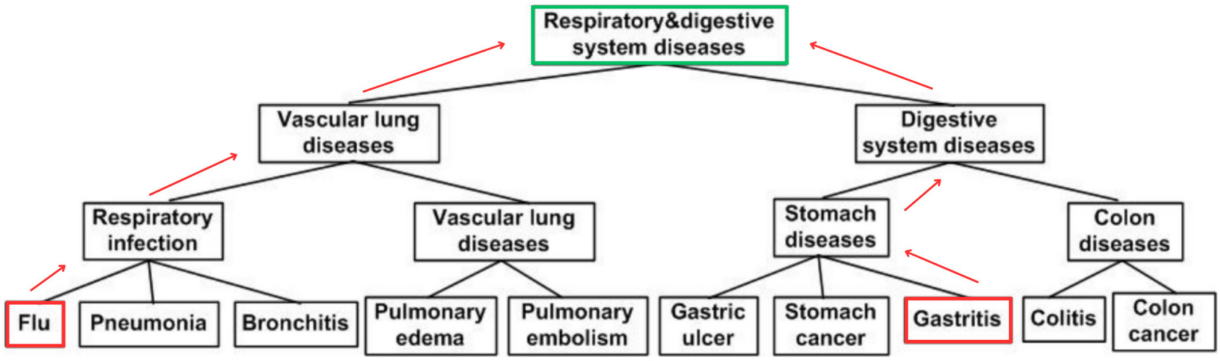
Ninghui Li, Tiancheng Li e Suresh Venkatasubramanian presentano il seguente schema gerarchico contenente alcune delle patologie inserite nel dataset.



[10]

Al fine di calcolare la distanza tra due elementi x e y. è necessario determinare il primo nodo comune, chiamato "lowest common ancestor" (LCA), dividendo il numero di "passi" effettuati per raggiungerlo con l'altezza dell'albero fino al livello immediatamente prima delle foglie (ossia 3).

[9][10] Li, Ninghui; Li, Tiancheng; Venkatasubramanian, Suresh (2007). 5.3 Analysis of t-Closeness with EMD, "T-Closeness: Privacy Beyond k-Anonymity and l-Diversity". t-Closeness: Privacy beyond k-anonymity and l-diversity



Si supponga di voler determinare la distanza tra "Influenza" e "Gastrite": il numero di passi effettuati è pari a 3 che, diviso per la profondità dell'albero come prima definita pari anch'essa a 3, indica il valore 1 (massima distanza possibile)

Conseguentemente, è facile determinare ogni altra distanza e comprendere come un valore inferiore rappresenti una maggior "vicinanza" (o, appunto, *closeness*) tra i dati.

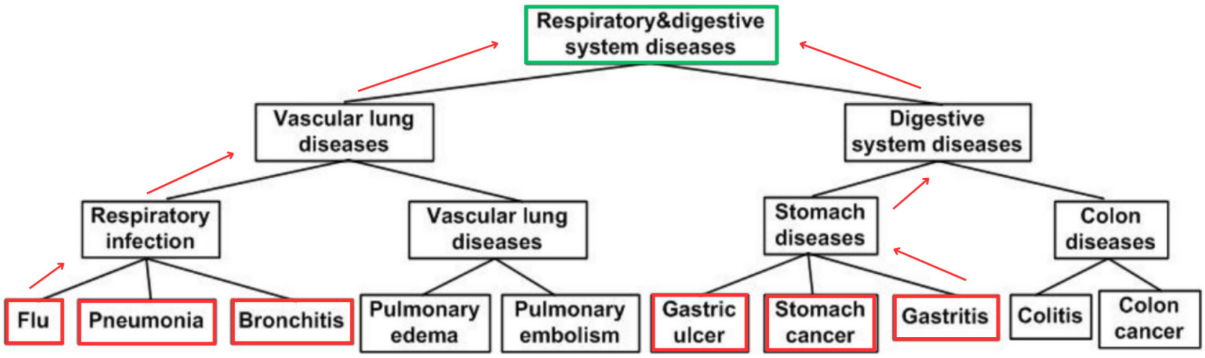
Alla luce di quanto esposto, com'è possibile quindi soddisfare la proprietà con il dataset riportato come esempio?

Si considerino singolarmente le classi di equivalenza

	Patologia	
Cl. eq. 1	Ulcera gastrica	
	Gastrite	Classe 1: {Ulcera gastrica, Gastrite, Carcinoma gastrico}
	Carcinoma gastrico	
Cl. eq. 2	Gastrite	
	Influenza	Classe 2: {Gastrite, Influenza, Bronchite}
	Bronchite	
Cl. eq. 3	Bronchite	
	Polmonite	Classe 3: {Bronchite, Polmonite, Carcinoma gastrico}
	Carcinoma gastrico	

La distribuzione originale è definita come:
{Ulcera gastrica, Gastrite, Carcinoma gastrico, Influenza, Polmonite, Bronchite}

Al fine di calcolare la distanza tra la la distribuzione originale e la classe di equivalenza n.1, è necessario considerare i 3 oggetti in quest'ultima contenuti:

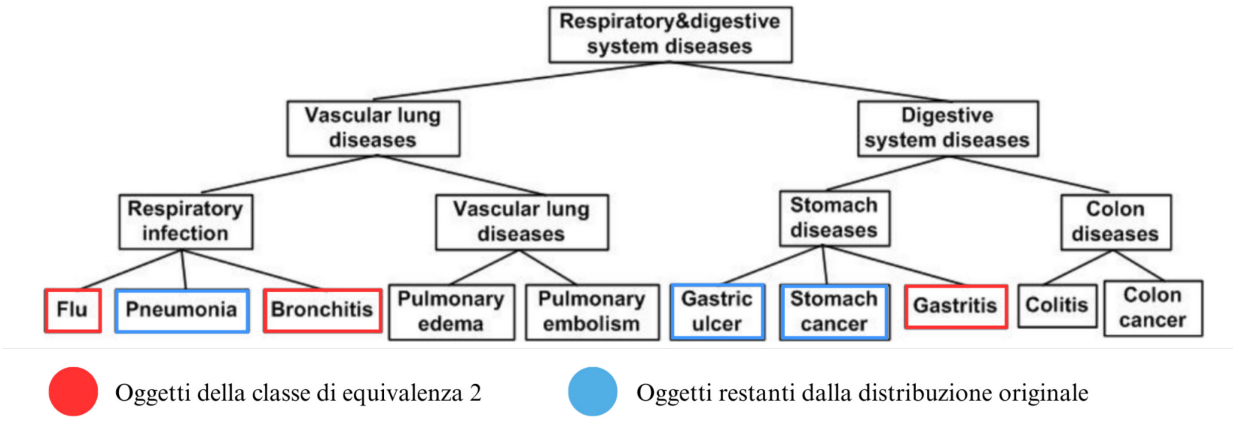


Anche nella presente analisi, la *EMD* è pari ad 1 per tutti gli oggetti della distribuzione della classe di equivalenza 1.

Analogamente al caso numerico, è necessario sommare le distanze di ciascun oggetto e dividere il risultato per la numerosità della distribuzione originale, quindi:

(1 + 1 + 1) / 6 = 0.5

Si valuti ora la seconda classe di equivalenza.



Il confronto oggetto per oggetto può essere arbitrariamente eseguito. Per il presente esempio, si valuteranno le seguenti distanze:

- Da Influenza a Polmonite: un solo passo per raggiungere il nodo *LCA*. Distanza: **1/3**
- Da Bronchite ad Ulcera gastrica: 3 passi per l'*LCA*. Distanza: 3/3=**1**
- Da Carcinoma gastrico a Gastrite: 1 passo per l'*LCA*. Distanza: **1/3**

(1/3 + 1 + 1/3) / 6 = 0.278 → Valore più prossimo alla distribuzione originale → maggior liv
di protezione dall'ottenimento di dati sensibili.

In sostanza, dai seguenti valori è possibile comprendere in che modo convenga riorganizzare un dataset al fine di ottenere il massimo livello di anonimizzazione possibile. È altresì vero che calcoli come quelli illustrati risultino eccessivamente prolissi qualora il database presentasse un'alta numerosità, e che in taluni casi possa non esistere una singola soluzione ottimale.

Nell'esempio qui presentato, una possibile riorganizzazione del dataset può essere la seguente:

	Età	CAP	Reddito	Patologia
Cl. eq. 1	<= 40	4767*	3.000	Ulcera gastrica
	<= 40	4767*	5.000	Carcinoma gastrico
	<= 40	4767*	9.000	Polmonite
Cl. eq. 2	> 40	4790*	6.000	Gastrite
	> 40	4790*	11.000	Influenza
	> 40	4790*	8.000	Bronchite
Cl. eq. 3	<= 40	4760*	4.000	Gastrite
	<= 40	4760*	7.000	Bronchite
	<= 40	4760*	10.000	Carcinoma gastrico

Dataset riorganizzato

La classe di equivalenza 2 non ha subito variazioni in quanto il valore della *EMD* era sufficientemente basso da non richiedere alterazioni. Le altre due classi, tuttavia, sono state riorganizzate.

Si può immediatamente notare come i range di età siano stati modificati introducendo semplicemente un limite, destro o sinistro, pari a 40 e come il mascheramento dei cap sia stato limitato all'ultima cifra: grazie a tali alterazioni, è stato possibile riorganizzare i record in modo che in ciascun gruppo la distribuzione dei redditi sia la medesima, così come quella delle patologie.

In altre parole, se volessimo identificare i dati sensibili di Bob (22 anni, CAP 47602), riusciremmo a definire il suo gruppo di equivalenza (il terzo), ma non saremmo in grado di comprendere qualora il suo reddito sia alto o basso, o di che genere di patologia soffra (respiratoria o gastrica?).

Il dataset qui illustrato soddisfa quindi tutte 3 le proprietà descritte:

- k-anonymity: il dataset è *3-anonymous* in quanto esistono almeno 3 diversi record per ciascuna classe di equivalenza;
- l-diversity: il dataset è *3-diverse* in quanto sono presenti almeno 3 attributi sensibili differenti in ciascuna classe di equivalenza;
- t-closeness: il dataset presenta *0.167-closeness* nell'attributo "Stipendio" e *0.278-closeness* nell'attributo "Patologia".

È bene evidenziare come tali proprietà non siano le uniche presenti nell'attuale stato dell'arte, in quanto tecniche derivate (per citarne alcune, k^m-anonimity o multi-dimensional k-anonimity) o nuove vengono spesso sviluppate in un contesto in continua evoluzione come quello tecnologico. Le 3 descritte tuttavia sono indubbiamente tra le più affermate.

Differential Privacy

Le tecniche più diffuse di anonimizzazione dei dati personali risultano essere relativamente semplici da implementare in maniera automatizzata, ma non garantiscono una matematica certezza per quanto concerne la privacy: infatti, attacchi che sfruttano tecniche di *data linkage* mediante analisi incrociate con altri dataset possono essere eseguiti con successo (v. studio di Latanya Sweeney o caso *Netflix Prize*[1]).

L'approccio definito *differential privacy*, al contrario, promette un livello di sicurezza basato su un *framework* matematico, in grado di garantire robustezza senza influire in modo eccessivamente significativo sull'utilità dei dati, specialmente in caso di analisi e studi statistici su larga scala.

Ad enunciarlo per la prima volta fu Cynthia Dwork, informatica esperta in cybersicurezza, nonché professoressa associata presso l'Università di Harvard e ricercatrice della multinazionale Microsoft.

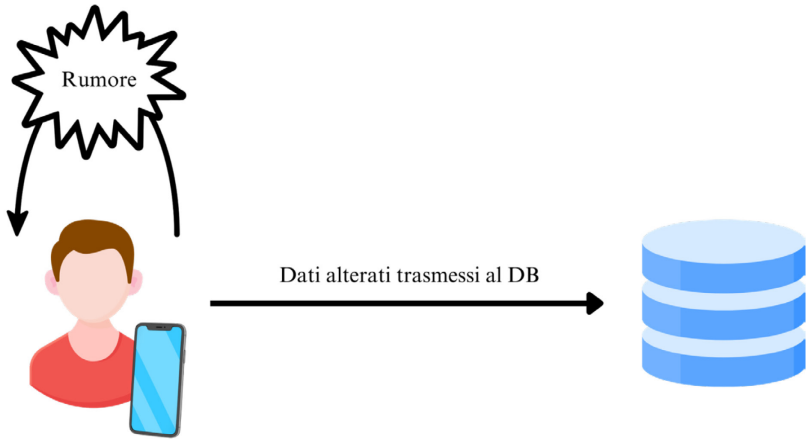
La differential privacy affronta il paradosso di entrare a conoscenza di informazioni utili riguardanti la popolazione, senza apprendere alcun dato relativo ai singoli individui; in altre parole, l'impatto che l'individuo può potenzialmente subire a seguito dell'analisi delle proprie informazioni non è superiore a quello che subirebbe se le sue informazioni non fossero proprio usate. Per chiarire il concetto, Cynthia Dwork presenta un esempio decisamente convincente: la consapevolezza che il fumo di sigaretta causi il cancro, a seguito di uno studio medico, può portare ad un aumento dei premi assicurativi relativi a spese mediche di lungo periodo (assumendo che la compagnia assicurativa sia a conoscenza delle abitudini da fumatore del proprio cliente). Si supponga che i dati relativi alle abitudini del suddetto cliente non siano in alcun modo stati trattati ai fini dello studio stesso, e nonostante ciò, ne debba ugualmente subire le conseguenze: infatti, nell'articolo della Dwork[2], si sostiene come siano le conclusioni di un'analisi ad influenzare l'individuo, non il fatto che egli sia stato o meno coinvolto; la conoscenza generale dell'informazione non può che coinvolgere l'intera popolazione di fumatori.

Le premesse di un sistema così robusto ed in grado di preservare la privacy degli individui mediante garanzie matematiche sono indubbiamente importanti, ma appaiono soddisfatte a seguito di un'attenta analisi del suo funzionamento.

Al fine di ottenere un dataset che soddisfi le *differential privacy*, del *rumore* viene aggiunto ai dati degli individui.
La perturbazione delle informazioni può avvenire **localmente** o in modo **centralizzato**.
Nel primo caso, i dati degli individui vengono alterati a livello *client*, prima dell'effettiva trasmissione al database.

---

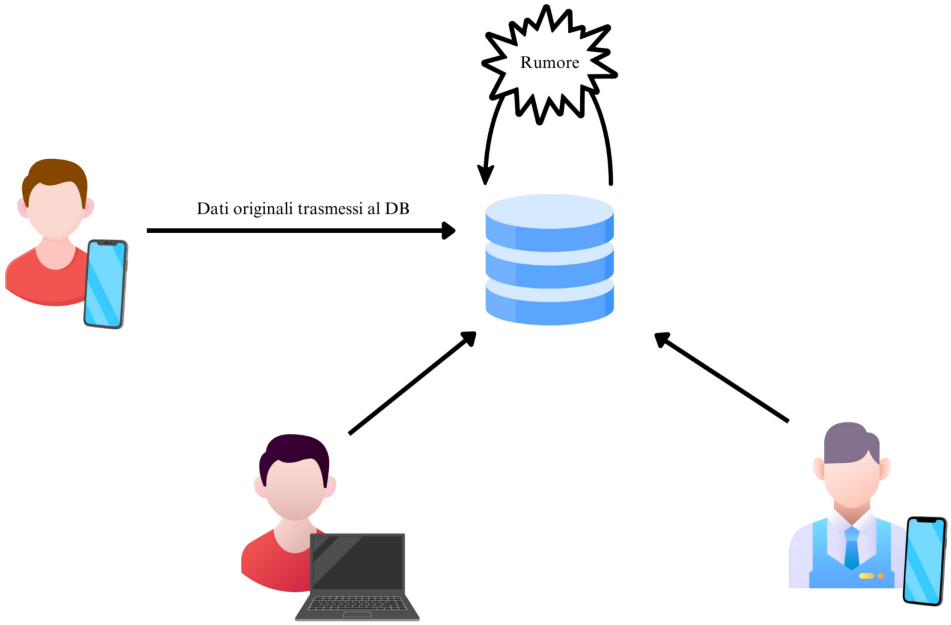
[1] Arvind Narayanan, Vitaly Shmatikov, "Robust De-anonymization of Large Sparse Datasets" - The University of Texas at Austin.
Nel 2006, Netflix ha rilasciato al pubblico un database con più di 100 milioni di recensioni, complete di date di pubblicazione, relative a circa 500.000 utenti. La finalità era organizzare un contest aperto per sviluppare un algoritmo di raccomandazione dei titoli caratterizzato da un'accuratezza superiore al 10%, valore raggiunto dall'allora algoritmo di Netflix.
Nonostante qualsiasi dato personale fosse stato omesso, Arvind Narayanan e Vitaly Shmatikov furono in grado di re-identificare numerosi utenti tramite un'analisi incrociata con i dati pubblicamente disponibili su IMDb (diversi utenti hanno rilasciato recensioni utilizzando il proprio nome).
[2] The Algorithmic Foundations of Differential Privacy, da Foundations and Trends R in Theoretical Computer Science Vol. 9, Nos. 3–4 (2014) 211–407 c 2014 C. Dwork and A. Roth DOI: 10.1561/0400000042



Il seguente approccio presenta un grande vantaggio, ovverosia la sicurezza che le informazioni originali non vengano in alcun modo trasmesse a terzi. Ciò è particolarmente utile nei casi in cui il titolare del dataset non risulti affidabile, o comunque per proteggersi da eventuali *data breach*. D'altra parte, il rumore applicato lato client risulta essere più permeante e, di conseguenza, i dati subiscono un'alterazione maggiore nella loro utilità.

L'aggiunta locale di *rumore* avviene sulla base di un *worst-case scenario*, trattando i dati come se presentassero il massimo livello di rischio per la privacy dell'individuo: l'offuscamento è, di conseguenza, maggiore. Va sottolineato anche che il livello di rumore sia proporzionale al numero di individui coinvolti, in quanto ogni dato presenta la propria componente di rumore proveniente da molteplici fonti (il rumore viene aggiunto da ogni partecipante in modo indipendente, l'effetto cumulativo di ciò può essere nettamente maggiore rispetto ad un modello centralizzato).

In contrapposizione al modello sopra descritto, l'approccio **centralizzato** mira maggiormente al mantenimento dell'utilità dei dati, limitando l'impatto del rumore, pur offrendo ottimali livelli di protezione per gli individui.



I dati vengono raccolti "in originale" tramite un database centralizzato ed il rumore viene aggiunto una sola volta alle informazioni aggregate. I dati presentano un livello di perturbazione inferiore e mantengono una maggiore utilità ai fini analitici; tuttavia, la trasmissione delle informazioni da client a server può essere compromessa, così come l'archiviazione centralizzata di queste. È a tal proposito

fondamentale implementare ulteriori misure di sicurezza, le quali possono esplicarsi in connessioni sicure e crittografate ai fini dell'invio, e la cancellazione dei dati originali dal server dopo l'applicazione del rumore.

A questo punto, definire il concetto di *rumore* e il modo in cui applicarlo è essenziale per comprendere il corretto funzionamento di una metodologia affermata come la *differential privacy*.

La perturbazione si basa sui principi matematici della *distribuzione di Laplace* (c.d. anche "doppia esponenziale"), una funzione di distribuzione probabilistica continua, contenente il parametro ε.

ε rappresenta in poche parole, secondo la definizione formale di *differential privacy*, il livello di protezione della privacy offerto.

Ad un valore di ε basso corrisponde un alto livello di privacy ed allo stesso tempo la necessità di introduzione di rumore più consistente, intaccando in parte l'utilità dei dati.

Il tutto può essere definito mediante la seguente disuguaglianza:

$$\Pr[M(x) \in S] \leq e^{\varepsilon} \Pr[M(y) \in S]$$

x ed y rappresentano due *neighboring datasets*, ossia dataset che differiscono per un elemento (si supponga che *x* contenga i dati di tutti i residenti di una città ed *y* ne contenga tutti meno uno).

La disequazione si legge: la probabilità che un meccanismo *M* (come una query in cui si richiede una media dei valori di un dataset) applicato ad *x* produca un certo output *S* è approssimativamente la stessa rispetto al caso in cui *M* fosse applicato ad *y*, il cui livello di similitudine con *x* dipende dal parametro ε (scritto e^{ε} per una maggior facilitazione nel calcolo grazie all'utilizzo di logaritmi, il che oltrepassa lo scopo della presente definizione).

La conseguenza pratica della formula è che il risultato delle analisi dei dataset è sostanzialmente lo stesso, a prescindere dal fatto che un individuo sia o meno presente in uno dei due[3].

È interessante conoscere come il concetto su cui si basa la DP, ossia il "conoscere informazioni su un gruppo di individui senza conoscere alcuno di questi", fosse stato già teorizzato da Stanley L. Warner nel 1965 tramite il meccanismo di *randomized response*[4].

Si è notato come gli utenti tendano a non comunicare il vero quando viene loro posta una domanda in merito ad informazioni sensibili: si immagini un sondaggio finalizzato a comprendere che percentuale di individui di una determinata città sia coinvolta in attività illecite; sarebbe chiaramente contro l'interesse degli interessati rispondere affermativamente.

A tal proposito, si propone di lanciare segretamente una moneta non truccata. Se al primo lancio esce testa, il partecipante deve rispondere onestamente alla domanda; nel caso di croce, è necessario

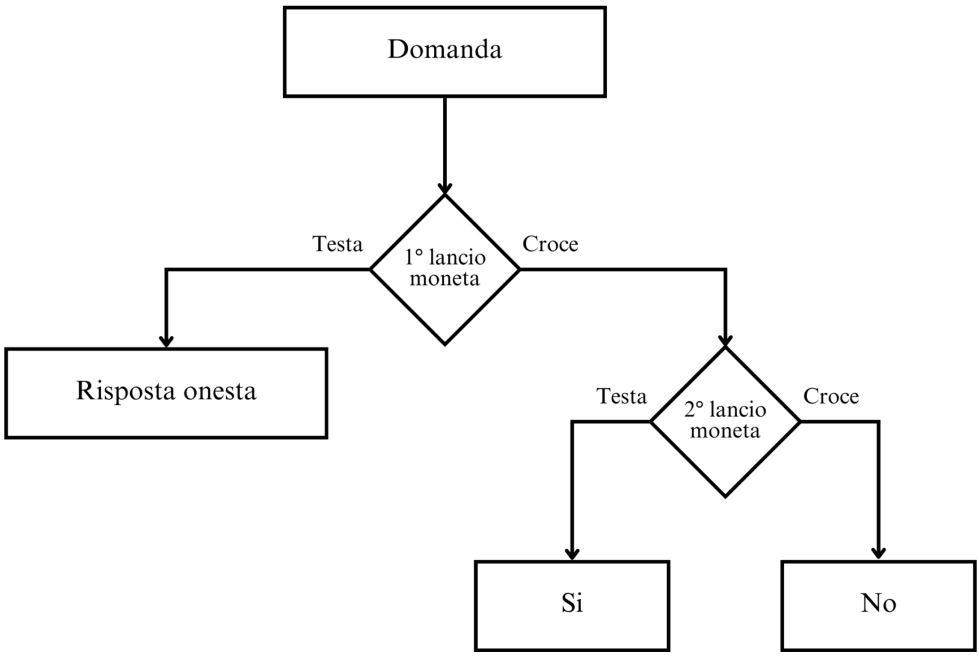
---

[3] "Specifically, it ensures that any sequence of outputs (responses to queries) is "essentially" equally likely to occur, independent of the presence or absence of any individual." Da "The Algorithmic Foundations of Differential Privacy, da Foundations and Trends R in Theoretical Computer Science" Vol. 9, Nos. 3–4 (2014) 211–407 c 2014 C. Dwork and A. Roth DOI: 10.1561/0400000042, pag. 6

[4] Warner, S. L. (March 1965). "Randomised response: a survey technique for eliminating evasive answer bias". Journal of the American Statistical Association. 60 (309). Taylor & Francis: 63–69. doi:10.1080/01621459.1965.10480775. JSTOR 2283137. PMID 12261830. S2CID 35435339.

lanciare nuovamente la moneta. Se al secondo lancio il risultato è testa, il partecipante è tenuto a rispondere affermativamente al quesito; al contrario, negativamente.

Schematizzo il tutto per una più immediata comprensione.



Mediante semplici calcoli, è possibile definire la probabilità di ciascuna risposta:

- 50% per "risposta onesta", in quanto testa e croce al primo lancio sono eventi equiprobabili;
- 25% per "Si" e 25% per "No" al secondo lancio:
  La probabilità che venga lanciata una seconda volta la moneta è pari alla probabilità che esca croce al primo lancio, ovvero 0,5; La probabilità che esca, ad esempio, testa al secondo lancio è ancora 0,5. Calcolando la probabilità combinata 0,5 * 0,5 = 0,25, troviamo le probabilità delle risposte "Si" e "No".

Assumendo comunque che tutti gli individui non coinvolti in attività illecite rispondano onestamente, è possibile comprendere come esista almeno il 25% di possibilità che una qualsiasi risposta sia non sincera: tale percentuale costituisce una negabilità plausibile che non rende possibile verificare la veridicità della risposta dei singoli individui, garantendo quindi un adeguato livello di sicurezza.
È lecito domandarsi se un *rumore* pari al 25% dei dati distorca in modo eccessivo le informazioni, inficiando i risultati di un'eventuale analisi: il vantaggio del presente metodo, tuttavia, è proprio conoscere il grado del rumore, in quanto è possibile effettuare una compensazione *ex-post* rettificando il 25% dei dati e ottenere un dataset sufficientemente accurato.

Il tema della *differential privacy* è estremamente vasto e complesso, tanto da meritare un approfondimento decisamente più ampio e dettagliato che, ai fini del presente elaborato, non è sfortunatamente possibile. È mia premura tuttavia indicare nella sezione **Bibliografia** l'insieme di *papers* e pubblicazioni utilizzate e schematizzate.

Capitolo 4: Case Study e Applicazioni Pratiche

1.Le Pronunce del Garante per la Protezione dei Dati Personali (GPDP)

Numerose sono le casistiche emerse nei recenti anni, menzionandone due:

- Provvedimento 1° giugno 2023, n. 226, comunemente denominato "Caso THIN" in merito all'utilizzo di *Real World Data* per ricerche nel settore medico. Il Garante ha sanzionato la società in quanto il processo di anonimizzazione dei dati personali mediante hashing del codice identificativo dei pazienti, oltre alla rimozione dai dataset degli attributi identificativi e la generalizzazione di età e residenza, non è stato ritenuto idoneo a minimizzare in modo adeguato il rischio di re-identificazione. L'applicazione del principio di k-anonimity non è avvenuta correttamente secondo il Garante, in quanto *"perde efficacia laddove, come nel caso in esame, a ciascun individuo sia associato un hash univoco (codice crittografico) seppur reso più complesso dalla presenza di un elemento di disturbo ignoto[1]".*
- Provvedimento del 18 luglio 2023, n. 311 (Caso Autorità di sistema portuale del Mare Adriatico settentrionale-Porti di Venezia e Chioggia). La pubblicazione di un reclamo sul sito istituzionale dell'Autorità, omettendo il solo nominativo del reclamante, non è stata ritenuta "*conforme alla disciplina rilevante in materia di protezione dei dati personali contenuta nel RGPD[2]"* in quanto risultavano presenti riferimenti a documenti e provvedimenti mediante i quali era possibile risalire univocamente all'identità dello stesso.

I casi sopra citati sottolineano come, per garantire la privacy degli individui, non si possa prescindere da uno studio accurato del contesto di applicazione delle misure tecniche. È bene tuttavia sottolineare come l'anonimizzazione dei dati personali non rappresenti un processo applicabile esclusivamente a dati trattati in maniera "testuale", bensì a qualsiasi genere di formato in grado di includere informazioni personali, come i file multimediali, trattati nel *case study* di seguito.

Videosorveglianza ed insufficienza nelle misure di anonimizzazione

Di particolare rilevanza è indubbiamente il Provvedimento dell'11 gennaio 2024, n.5 del Garante per la Protezione dei Dati Personali, con il quale l'Autorità ha sanzionato il Comune di Trento per l'illecito trattamento di dati personali e particolari degli individui mediante sistemi di sorveglianza.

La fattispecie in oggetto ha visto il Comune di Trento partner di tre progetti finanziati dall'UE, quali programmi di ricerca MARVEL, PROTECTOR e PRECRISIS.

In particolare, il progetto MARVEL (*"Multimodal Extreme Scale Data Analysis for Smart Cities Environments"*) ha reso possibile l'acquisizione di materiale sia videografico, mediante 14 videocamere di sorveglianza già installate nelle pertinenze comunali ed utilizzate ai fini di sicurezza urbana, che audio, mediante 6 microfoni posizionati ai fini del progetto stesso.
Il Comune sosteneva che ogni dato raccolto dai suddetti dispositivi sarebbe stato anonimizzato immediatamente a seguito della sua raccolta e analizzato mediante sistemi di intelligenza artificiale per identificare eventuali situazioni di pericolo per la pubblica sicurezza.

---

[1] Da Provvedimento 1° giugno 2023, n. 226, Garante per la Protezione dei Dati Personali
[2] Da https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9920562

L'accesso ai dati sarebbe stato possibile anche da parte della Fondazione partner, nonché fornitore delle infrastrutture hardware per il processamento e l'anonimizzazione dei dati, mediante VPN (rete privata virtuale o *Virtual Private Network*).

PROTECTOR (*"PROTECTing places of wORship"*), in aggiunta all'acquisizione di materiale videografico privo di tracce audio, prevedeva l'analisi di potenziali contenuti d'odio di carattere religioso su piattaforme social quali "Twitter" (ora "X") e "YouTube".
L'elaborazione dei dati avveniva tramite algoritmi di A.I. ed erano previsti, in particolare, per le informazioni visive:

- Object detection: tecnologia grazie alla quale è possibile identificare automaticamente la tipologia degli oggetti inquadrati (veicolo, pedone, cicli…), senza analizzarne la specifica identità;
- Visual tracking: sistema open source in grado di tracciare il movimento dei soggetti;
- Anomaly detection: a seguito di training dell'algoritmo tramite librerie pubbliche, il sistema di A.I. avrebbe permesso di identificare potenziali situazioni di pericolo. Le immagini analizzate sarebbero successivamente usate per il training stesso ai fini del miglioramento nell'accuratezza dei risultati.

L'anonimizzazione dei dati visivi prevedeva l'utilizzo di sfocatura di volti e targhe dei veicoli, mentre per l'audio era prevista un'alterazione delle caratteristiche identificative della voce.

Per quanto concerne l'ambito social network:

- Rilevamento di messaggi d'odio mediante analisi del contenuto testuale;
- Analisi delle emozioni nei post ai fini di rilevare, mediante dati aggregati, eventuali picchi d'ira o di aggressività;
- Rilevamento della disinformazione: tramite un'analisi semantica del linguaggio, PROTECTOR era in grado di identificare presunte fake news relative all'ambito religioso.

In merito a PRECRISIS, il progetto risultava ancora in fase di attivazione ed era privo di qualsiasi componente di intelligenza artificiale.

Il Comune di Trento ha successivamente dichiarato che ogni dato era stato acquisito mediante valida base giuridica, riconducibile all'art. 6 c. 7 della legge 38/2009 che recita che *"per la tutela della sicurezza urbana, i comuni possono utilizzare sistemi di videosorveglianza in luoghi pubblici o aperti al pubblico"* e alle disposizioni della Direttiva 2016/680, del D.lgs 51/2018 (attuazione della direttiva *"relativa alla protezione delle persone fisiche con riguardo al trattamento dei dati personali [...] a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali"*) e del D.l. 20 febbraio 2017, n.14 (*Disposizioni urgenti in materia di sicurezza delle città*).

Il Comune stesso ha dichiarato che i dati personali erano stati "anonimizzati alla fonte" e non utilizzati a fini di profilazione o condivisi a terze parti. La Fondazione stessa è stata identificata quale unico soggetto esterno al Comune a cui i dati sarebbero stati condivisi, tuttavia in forma già anonimizzata. Le informazioni sarebbero state utilizzate, sempre in forma anonima, per il training degli algoritmi di riconoscimento di situazioni di pericolo.

L'eventuale identificazione di situazioni anomale sarebbe stata segnalata alle autorità locali, monitorata poi nello specifico da operatori della Polizia Locale ai fini di valutare l'effettività del pericolo.

In seguito all'attività istruttoria da parte del Garante, si è rilevato come il Comune abbia posto in essere un trattamento di dati personali relativi a reati, in quanto l'utilizzo di videocamere, microfoni

e l'analisi di contenuti d'odio su piattaforme social, mira alla rilevazione di minacce alla sicurezza pubblica con particolare riguardo ai luoghi di culto.

Si è ritenuta non accettabile la tesi del Comune secondo la quale la raccolta di immagini mediante videocamere di sorveglianza non configurerebbe un trattamento di dati personali relativi a reati, salvo questi non vengano successivamente utilizzati per l'accertamento di questi; peraltro, l'*EDPB* (Comitato europeo per la protezione dei dati personali) aveva già precedentemente chiarito come l'implementazione di sistemi di videosorveglianza non configuri necessariamente un trattamento di dati sensibili, salvo questi non vengano ricavati tramite il loro utilizzo. Ripercorrendo quanto descritto nel presente elaborato, l'utilizzo di informazioni relative a credi religiosi rientra nella categoria dei "dati particolari" (ex art. 9, GDPR), che prevede un divieto nel trattamento di questi, fatte salve specifiche circostanze.

Gli stessi contenuti testuali, ricavati tramite social network, possono potenzialmente *"rivelare le convinzioni religiose dei relativi autori o di terzi menzionati in detti messaggi"*.

Era peraltro chiaro al Comune stesso come il trattamento dei dati sopracitati avrebbe ingenerato una sensazione di invasività nella sfera privata dei cittadini; tuttavia, il DPO dell'Ente, interpellato ai fini dei presenti progetti, aveva fornito un parere positivo a riguardo.

In merito alle **tecniche di anonimizzazione** impiegate, il Garante ha sollevato diverse criticità secondo un approccio particolarmente restrittivo, in particolare:

- Il processo di anonimizzazione dei dati, avvenuto in maniera automatica immediatamente dopo l'acquisizione delle immagini e dei contenuti audio, non può sottrarre il trattamento all'ambito di applicazione del GDPR: infatti, nonostante intercorra un lasso di tempo minimo tra l'acquisizione delle immagini e la loro anonimizzazione, non è sufficiente per sostenere la tesi difensiva del Comune.
- Il processo stesso di anonimizzazione non è stato ritenuto idoneo. Con particolare riguardo ai dati **audio**: la mera sostituzione della voce non altera in alcun modo il contenuto delle tracce, le quali possono rivelare dati personali del parlante o di terzi menzionati o interlocutori. Non è inoltre da escludersi come le conversazioni possano coinvolgere informazioni relative a soggetti fragili o vulnerabili come minori.
  Tra le tesi difensive presentate dal Comune di Trento appariva la limitata sensibilità dei microfoni che, secondo l'Ente, non sarebbero stati in grado di captare conversazioni ma unicamente suoni intensi, visto il posizionamento in altezza di diversi metri rispetto alla superficie. Vero che, come sostiene il Garante, il progetto MARVEL *"prevedeva, infatti, espressamente la raccolta dell'audio delle conversazioni intercorse nella pubblica via"*, il che implicitamente considerava possibile la registrazione delle conversazioni stesse, vista anche l'implementazione di tecniche di alterazione della voce dei soggetti parlanti.

  Per quanto concerne il materiale **video**, l'utilizzo di sfocature gaussiane e l'offuscamento degli identificatori come volti e targhe dei veicoli sono stati ritenuti inidonei a scongiurare il rischio di re-identificazione: infatti, a detta del Garante, i soggetti sarebbero stati potenzialmente identificabili mediante altre caratteristiche, peculiari per ciascun individuo, quali *"corporatura, abbigliamento, posizione nella scena filmata, caratteristiche fisiche particolari"*, informazioni detenute da terzi, o addirittura tramite percorsi e spostamenti effettuati.

  Le videocamere erano, inoltre, posizionate ad un'altezza variabile dai 3,5 m ai 40 m, in grado di produrre immagini con una risoluzione di 1200x1600 pixel, con *"elevata compressione delle immagini, che genera […] una alterazione dei dettagli"*, inducendo peraltro una

"distorsione prospettica" che avrebbe reso concretamente non possibile il riconoscimento di caratteristiche sufficienti per l'identificazione di soggetti, nonché di potenziali situazioni di rischio.
L'Autorità Garante sostiene pertanto che l'utilizzo di infrastrutture hardware incapaci di produrre immagini sufficientemente dettagliate sia incompatibile con la natura del progetto PROTECTOR.

Con riguardo ai contenuti **testuali** sulle piattaforme YouTube e Twitter, nella prima i nomi utenti venivano cancellati, mentre nella seconda solo pseudonimizzati mediante un ID casuale generato automaticamente.

Il Garante non ha poi ritenuto valida la base giuridica per la raccolta di informazioni in quanto l'ingerenza effettuata da parte del Comune di Trento nei confronti dei diritti degli individui (in particolare il diritto al rispetto della vita privata e il diritto alla protezione dei dati di carattere personale sanciti dalla Carta dei Diritti Fondamentali dell'UE) non era proporzionale agli obiettivi di protezione della pubblica sicurezza sostenuto dall'Ente, il quale non è stato peraltro in grado di fornire adeguate garanzie alla protezione dei dati dei soggetti.

Infine, a garanzia dei principi di "*liceità, correttezza e trasparenza*", l'Autorità ha giudicato inidonea l'informativa presentata agli interessati in quanto fraintendibile nelle finalità del trattamento ed incompleta in merito alle tipologie di registrazione audio, non specificando che il contenuto stesso delle conversazioni sarebbe stato potenzialmente registrato (nonostante il Comune si sia opposto a tale considerazione in quanto, come sopra menzionato, ritiene nulla la possibilità di registrazione di contenuti semantici per caratteristiche stesse della strumentazione). La raccolta di materiale audio non era tuttavia da considerarsi consentita per finalità di sicurezza urbana.
Diverse omissioni sono state rilevate anche per quanto concerne il trattamento dei dati relativi alle piattaforme social (delle quali manca completamente la menzione), così come per il trasferimento delle informazioni ai partner dei progetti, sul presupposto che il Comune avesse effettivamente anonimizzato i dati, rendendo quindi non obbligatoria l'informativa agli interessati.

Assente anche un'adeguata valutazione d'impatto dei trattamenti effettuati, in quanto la documentazione presentata dal Comune non risultava sottoscritta, né comprendente di analisi in merito alla "necessità e proporzionalità dei trattamenti in relazione alle finalità" o a rischi per diritti e libertà degli individui, limitandosi ad una mera considerazione sulle minacce alla sicurezza dei database.

Alla luce di quanto sopra e delle circostanze attenuanti, compresa la buona fede del Comune nell'incorrere in un errore in diritto, il Garante ha ordinato il pagamento della somma di €50.000 nonché l'interruzione del trattamento dei dati e la cancellazione di quanto già raccolto.

Risulta lecito a questo punto domandarsi, al netto delle inadeguatezze di informativa e valutazione d'impatto: in che modo il Comune di Trento avrebbe dovuto agire per evitare di incorrere in sanzioni in merito alle tecniche di anonimizzazione poste in essere?

L'Autorità Garante ha adottato un approccio particolarmente restrittivo in merito al trattamento di dati delle persone fisiche, con particolare riguardo ai materiali video: la presenza di fotogrammi non chiari in quanto di bassa definizione non è da ritenersi, a mio avviso, incompatibile con le finalità dei progetti. La scarsa quantità di dettagli impedisce il riconoscimento di tratti distintivi del volto degli individui, così come l'uso di sfocature previene l'identificazione diretta delle persone, ma non necessariamente rende impossibile l'analisi di situazioni anomale, in quanto potrebbero richiedere un livello di dettaglio inferiore per essere adeguatamente identificate (così come per i sistemi di

riconoscimento di oggetti).

Tale considerazione si tratta, tuttavia, di una mera ipotesi in quanto non è dato esaminare personalmente i fotogrammi tratti dalle videocamere: è tuttavia plausibile che, vista la risoluzione pari a 1600x1200px, la quale si traduce in 1.920.000 pixel totali (meno dell'8% di pixel di differenza rispetto alla comune risoluzione FullHD), sia ugualmente possibile riconoscere una potenziale situazione di pericolo da parte di un algoritmo. Vero è che, ai fini di valutare il livello di dettaglio di una ripresa, devono essere considerati numerosi ulteriori fattori come la tipologia di sensore che cattura le immagini, il tipo di lente, eventuali difetti ottici (aberrazioni, distorsioni da grandangolo, flaring, glaring, più comunemente "aloni di luce") e la compressione dei fotogrammi che può portare ad artefatti.

È tuttavia da chiedersi: se l'algoritmo di intelligenza artificiale è in grado di riconoscere una targa di un autoveicolo e quindi poterla offuscare, così come un volto, è davvero incapace di riconoscere situazioni anomale (assumendo che il riconoscimento derivi da una valutazione di particolari movimenti, assembramenti, posizioni dei soggetti inquadrati)?

Per quanto concerne l'ambito di anonimizzazione dei dati, il Garante sostiene che non solo gli identificatori comunemente riconosciuti come tali siano in grado di portare al riconoscimento di un individuo (e.g. volto, targhe di veicoli), ma anche *"corporatura, abbigliamento, posizione nella scena filmata, caratteristiche fisiche particolari"* e *"informazioni relative al percorso effettuato da una determinata persona individuata nelle immagini video mediante le predette caratteristiche fisiche e gli elementi di contesto"*.

Riprendiamo il concetto di "dato personale", citando l'articolo 4, comma 1, del GDPR:
«dato personale»: qualsiasi informazione riguardante una persona fisica identificata o identificabile («interessato»); si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come il nome, un numero di identificazione, dati relativi all'ubicazione, un identificativo online o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale;

È chiaro come il Garante sostenga che, contestualizzando le situazioni in cui un determinato soggetto agisce o nelle quali la sua immagine viene catturata, il rischio di identificazione non sia assolutamente assente; ciò porta, tuttavia, ad un quesito: com'è possibile anonimizzare una sequenza di fotogrammi provenienti da videocamere di sorveglianza? O addirittura, è di fatto possibile anonimizzarla?

Un'opzione di anonimizzazione può essere rappresentata dall'offuscamento non solo del volto degli individui, ma del contesto intero: si pensi alla situazione in cui alcuni individui vengono ritratti all'interno di una via; offuscando i volti e il resto dell'ambiente, pur mantenendo in nitidezza la fisionomia delle persone stesse, si potrebbe diminuire il rischio di identificazione con riguardo al *contesto*. Permangono tuttavia diversi dubbi: se si è a conoscenza dell'esatta videocamera che ha registrato una sequenza di immagini, è di immediata comprensione anche il luogo in cui i fatti avvengono. È inoltre possibile che alcune vie particolarmente frequentate presentino una quantità tale di individui ripresi da non rendere efficace l'offuscamento del contesto, il quale sarebbe rappresentato da aree di immagine così minute da rendere irrilevante qualunque sfocatura. Tutto ciò, infine, non prende in considerazione la potenziale presenza di persone caratterizzate da una fisionomia riconoscibile, ossia dai sopracitati *elementi caratteristici* dell'identità fisica di una persona. Offuscare completamente l'individuo, d'altra parte, renderebbe vano ogni tentativo di analisi della situazione, minando la finalità stessa delle riprese.

Potrebbe essere esplorabile l'utilizzo di tecniche in grado di sfruttare modelli generativi di A.I. al fine

di alterare i fotogrammi ed introdurre dal vivo, ad esempio, cloni di individui presenti all'interno della scena in modo da duplicare persone già presenti e prevenire l'univoca identificazione delle stesse (alla stregua dei principi della *k-anonymity*). In altre parole, con due figure "clonate" può risultare più difficoltoso individuare quale delle due sia effettivamente coinvolta in una determinata situazione (ad es. il percorso effettuato). Non c'è tuttavia dubbio su come questa opzione richieda l'utilizzo di algoritmi e potenza di calcolo di altissimo livello e, comunque, solleverebbe questioni in merito alla sua accuratezza e a potenziali rischi legati a "falsi allarmi" (se gli individui automaticamente generati fossero erroneamente coinvolti in fittizie situazioni anomale, la loro presenza potrebbe portare ad interventi a vuoto da parte delle Autorità).

Una situazione simile porta inevitabilmente a discutere su come un'Intelligenza Artificiale più sviluppata della "semplice" *object detection* possa effettivamente risultare d'aiuto in situazioni simili; a tal proposito, il prossimo capitolo si concentrerà proprio sull'utilizzo dell'A.I. e sulle nuove frontiere normative rappresentate dall'A.I. Act.

Infine, la decisione del Garante in merito alla mancata anonimizzazione delle tracce **audio** è, a mio avviso, giusta: l'altezza dei microfoni può sicuramente portare ad un suono confuso ed inintelligibile, ma non è da escludersi che alcune informazioni personali si possano ugualmente captare. Ci si immagini la situazione in cui un individuo menziona a voce particolarmente alta una terza persona, in un momento di scarsa frequentazione della zona (e quindi con limitato rumore di sottofondo). È chiaro come si tratti di un contesto estremamente ristretto e che non rappresenta il comune svolgersi della vita quotidiana all'interno delle vie della città; è altresì vero che ciò non preclude il fatto che possa effettivamente realizzarsi. I microfoni quindi, secondo la mia opinione, presentano un rischio di raccolta di dati personali e la semplice sostituzione della voce, per quanto utile a proteggere la privacy del parlante, non esclude che il contenuto presenti informazioni personali altrui (o del parlante stesso).
A tal proposito, potrebbero risultare d'aiuto algoritmi di Intelligenza Artificiale per il riconoscimento di potenziali dati personali (nomi o altri identificativi), ma quantomeno all'attuale stato dell'arte non ritengo siano sufficienti, visto il grande numero di possibili informazioni personali menzionabili e la scarsa accuratezza in ambito di *speech recognition*, specialmente se applicata a segnali audio di scarsa qualità, per non parlare del fatto che comunque si tratterebbe di una modifica non realizzabile "dal vivo", ma solo a registrazione effettuata.

In conclusione, il Garante ha sicuramente applicato le normative vigenti con la massima attenzione nei confronti della protezione delle persone fisiche, creando però di fatto dubbi su come effettivamente sia possibile rendere anonimi un file video o una traccia audio. A mio avviso, al di là dell'ineccepibile decisione in merito all'inadeguatezza della valutazione d'impatto realizzata dal Comune, per quanto descritto in atti, è complicato immaginare come una qualsiasi organizzazione possa porre in essere progetti del genere viste le importanti limitazioni imposte dall'Autorità Garante. D'altra parte, è condivisibile l'approccio restrittivo alla luce delle potenziali ripercussioni sulla percezione di libertà dei cittadini.

Con una legislazione particolarmente stringente ed una visione altrettanto protettiva dell'individuo, in che modo è possibile garantire l'utilità del dato personale, pur ottemperando alle normative vigenti?

2.Le nuove frontiere della tecnologia

A.I. Act e dati sintetici

Il Regolamento n. 1689 del 13 giugno 2024, chiamato "A.I. Act", si propone come principale (nonché prima al mondo[1]) fonte normativa finalizzata a regolare l'utilizzo di sistemi di intelligenza artificiale.

Si tratta di un chiaro segnale di innovazione da parte dell'Unione Europea che, nel corso degli anni in ambito A.I., ha dimostrato di essere un passo indietro rispetto alle superpotenze di U.S.A. e Cina: la creazione di un framework robusto, *human-centric* e che mira a mitigare i rischi legati all'intelligenza artificiale secondo un approccio *by design*, può porre le basi ad un corretto e sano sviluppo delle nuove frontiere dell'intelligenza artificiale.
Infatti, l'A.I. Act sottolinea nel Considerando n.69 come i principi della privacy debbano essere soddisfatti durante l'intero ciclo di vita dell'intelligenza artificiale e che, oltre a tecniche quali l'anonimizzazione, vengano adottate ulteriori tecniche a garanzia della protezione dei dati personali.

Il regolamento suddivide gli algoritmi di intelligenza artificiale in categorie, sulla base dei rischi che li caratterizzano:

1. **Rischio minimo**: come filtri anti-spam, risultano generalmente esenti da obblighi e possono essere utilizzati senza particolari restrizioni. Resta ugualmente fortemente suggerita l'implementazione dei principi di *fairness*, *human oversight* e *non-discrimination*.
2. **Rischio limitato**: categoria in cui risulta obbligatorio informare l'utente del fatto che il suo "interlocutore" sia un algoritmo di intelligenza artificiale (in ottemperanza al principio di *trasparenza*), salvo non sia palese dal contesto d'utilizzo.
3. **Alto rischio**: algoritmi che, in caso di utilizzo scorretto o di risultati errati, possono causare gravi pregiudizi agli individui, con particolare riguardo ai diritti fondamentali come salute, sicurezza e libertà personali.
4. **Rischio inaccettabile**: algoritmi di A.I. incompatibili con i principi e le libertà fondamentali dell'Unione, quali sistemi di credito sociale, categorizzazione degli individui su base biometrica o di altre informazioni sensibili, manipolazione subliminale, polizia predittiva etc.

È chiaro come i principi del regolamento di leghino indissolubilmente a quelli del GDPR, adottando un approccio basato sul rischio e offrendo importanti livelli di protezione con riguardo ai dati personali e particolari. L'analisi di questi, tuttavia, può offrire numerosi vantaggi e, ancora una volta, ci si trova costretti a bilanciare utilità e sicurezza. In risposta al quesito posto al termine della precedente sezione, una promettente soluzione in grado di conciliare entrambi i principi appare essere l'uso di **dati sintetici**.
Un dato sintetico è un'informazione generata artificialmente, mediante l'utilizzo di algoritmi e modelli matematici, in grado di replicare situazioni reali. Utilizzati in particolar modo in contesti ove garantire la privacy degli individui o la raccolta di informazioni dal pubblico risultano complesse, i dati sintetici sfruttano dati personali già raccolti, analizzandoli mediante modelli di apprendimento automatico come GAN (Generative Adversarial Networks) e VAE (Variational AutoEncoders), metodi statistici come distribuzioni poissoniane o normali e simulazioni.

[1]https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

I dati generati a seguito dell'elaborazione tramite algoritmi di *machine learning* si presentano come repliche di dati "reali", privi però di riferimenti ad individui. Il processo non risulta nemmeno essere reversibile, in quanto dal dato sintetico non è possibile risalire all'esatta informazione reale.

La raccolta di informazioni personali risulta essere complessa ed onerosa (sia in termini economici che di tempistica), oltre a presentare chiari problemi di privacy: l'utilizzo di dati sintetici è, pertanto, una soluzione agile e comunque contemplata dall'A.I. Act. Il regolamento stesso, classificandoli come dati non personali[2], offre indirettamente un più ampio spazio di manovra ai soggetti che ne fanno utilizzo, non dovendo ottemperare agli obblighi del GDPR. Inoltre, all'articolo 10 del testo normativo ed in particolare nel paragrafo 5, lettera a), l'utilizzo dei dati sintetici viene implicitamente incoraggiato: nei casi in cui si preveda il trattamento di categorie di dati particolari, l'eliminazione di eventuali distorsioni idonee ad "avere un impatto negativo sui diritti fondamentali o di comportare discriminazioni"[3] dovrebbe essere effettuata mediante "dati sintetici o anonimizzati"; nel caso in cui i suddetti dati non fossero in grado di correggere tali criticità, ai fornitori di sistemi di I.A. ad alto rischio è data la possibilità di utilizzare dati particolari, fatte salve le garanzie alle lettere f) e g) del paragrafo 2.

Quanto appare come una soluzione in grado di coniugare un'ottimale qualità dell'informazione alla massima protezione della privacy non è, tuttavia, esente da svantaggi e significative limitazioni.

In primis, un dato sintetico si basa su dati realmente acquisiti, i quali devono chiaramente essere trattati nel rispetto delle normative vigenti; peraltro, la qualità delle informazioni, prodotte a seguito di elaborazioni mediante modelli di machine learning, dipende da quella dei dati forniti. Per tali ragioni, non risulta possibile definire i dati sintetici come intrinsecamente anonimi.

È inoltre utile introdurre il concetto di *fedeltà* delle informazioni generate artificialmente, ossia la capacità di queste di riprodurre accuratamente dati reali: in caso di eccessiva fedeltà, i dati possono presentare pattern o correlazioni tali da permettere la ricostruzione delle informazioni originali. Tale ricostruzione può non portare all'esatto dato originale vista la natura irreversibile del processo di sintetizzazione dei dati, ma può fornire un sufficiente livello di accuratezza al fine di re-identificare un individuo. Fedeltà e qualità dei dati in ingresso sono inoltre indissolubilmente legati al rischio di *bias* nelle informazioni prodotte: un dataset che presenta bias, riprodotto con un livello di fedeltà elevato, può trasmettere gli stessi problemi ai dati sintetici risultanti[4].

La creazione di dati sintetici è, inoltre, dipendente dall'algoritmo utilizzato. Date la natura "*black box"* dei sistemi di intelligenza artificiale e, conseguentemente, le limitate possibilità di comprensione del loro funzionamento (nonostante la crescente rilevanza del concetto di *explainable A.I.*), le metodologie di generazione dei dati sintetici non risultano trasparenti.

Appare inoltre non immediata la valutazione in merito all'accuratezza stessa dei dati sintetici, i quali potrebbero non rappresentare con sufficiente precisione situazioni reali: fattori come la presenza di *outliers* (dati con un'elevata varianza rispetto alla norma), in taluni casi rilevanti ai fini dell'analisi di un dataset, potrebbero non essere replicati in modo corretto.

---

[2] Regolamento 1689/2024, Art.59 par.1 lett.b): "[...] mediante il trattamento di dati anonimizzati, sintetici o di altri dati non personali"

[3] Regolamento 1689/2024, Art.10 par.5 lett.a): "il rilevamento e la correzione delle distorsioni non possono essere realizzati efficacemente mediante il trattamento di altri dati, compresi i dati sintetici o anonimizzati;"

[4] "For synthetic data to be meaningful, it must be similar to and different from the original data in some sense. If synthetic data is being considered, then there is a reason that the original data is inappropriate or inadequate for the task at hand – be it because it is non-private, biased, or too small – and so synthetic data that is too similar to the original data will also suffer from the same problems." tratto da "Synthetic Data - what, why and how?" arXiv:2205.03257 [cs.LG] – J. Jordon, L. Szpruch, F. Houssiau, M. Bottarelli, G. Cherubin, C. Maple, S. N. Cohen, A. Weller

Nonostante ciò, i dati sintetici presentano numerosi vantaggi che, spesso, superano le difficoltà sopra illustrate: contesti in cui la perfetta fedeltà dei dati sintetici rispetto a situazioni reali non è un requisito essenziale, come simulazioni in campo marketing, possono indubbiamente beneficiare dall'uso di questi. L'utilizzo di dati personali a scopo commerciale presenta numerose limitazioni normative, in particolar modo per quanto concerne il consenso alla raccolta ed al trattamento; a tal proposito, la possibilità di sfruttare dati in grado di replicare con sufficiente accuratezza dinamiche del mondo reale non può che risultare vantaggiosa.

Inoltre, l'uso di dati sintetici per finalità quali il riconoscimento di oggetti o di persone, videosorveglianza compresa, è indubbiamente un campo di applicazione da esplorare, sia dal punto di vista del *training* dell'algoritmo, che per quanto concerne l'effettivo utilizzo di questo.
Si pensi all'utilizzo di volti artificiosamente generati (c.d. *deepfake*, come regolati dall'A.I. Act) in procedure di *training* algoritmico: il sistema di A.I. sarebbe in grado di aumentare le proprie capacità di riconoscimento di volti, senza tuttavia sfruttare immagini di individui reali.

Per quanto concerne l'impiego del sistema stesso in una fattispecie analoga a quella relativa al Comune di Trento, è possibile adottare tecniche di anonimizzazione tramite la sostituzione dei volti reali con volti sintetici, così come altre caratteristiche identificative, pur mantenendo ogni proprietà utile ai fini della sicurezza pubblica (per esempio, dettagli come la presenza di un'arma nella mano di un individuo devono restare inalterati).

Entrambi i campi di applicazione presentano, anche in questo caso, importanti limitazioni: è essenziale che l'addestramento degli algoritmi si basi su dataset privi di *bias*; di conseguenza, è necessario che anche i *deepfake* utilizzati durante il training ne siano privi. In caso contrario, addestrare un algoritmo con dati provenienti da una precedente elaborazione e già caratterizzati da *bias*, non può che amplificare ed esacerbare tali inaccuratezze.
D'altra parte, anche la sostituzione di volti e di caratteristiche particolari degli individui ritratti nei filmati deve risultare priva di *bias*. Si pensi, ad esempio, alla situazione in cui i volti sintetici associati ad individui coinvolti in atti illeciti presentino tratti distintivi di una specifica etnia.

Non vi è dubbio su come i dati sintetici possano effettivamente portare numerosi vantaggi viste facilità d'utilizzo e limitate costrizioni in ambito privacy, ma considerarli la soluzione al problema del *trade-off* sicurezza/utilità è, a mio avviso, prematuro. Vero è che grandi aziende come Amazon sfruttano dati sintetici per il riconoscimento vocale negli *smart assistant*[5], American Express[6] e J.P. Morgan[7] a fini di riconoscimento di frodi e prevenzione di riciclaggio di denaro; tuttavia, valutare nel concreto il livello di protezione degli individui non risulta per nulla scontato e, in ogni caso, non è possibile prescindere dall'applicazione di ulteriori misure di sicurezza.

[5] https://www.gartner.com/en/newsroom/press-releases/2022-06-22-is-synthetic-data-the-future-of-ai
[6] https://blogs.nvidia.com/blog/american-express-deep-learning/
[7] https://www.jpmorgan.com/technology/technology-blog/synthetic-data-for-real-insights

Conclusioni

L'anonimizzazione dei dati personali rappresenta una pratica fondamentale nel contesto attuale, caratterizzato da un uso sempre più intensivo di informazioni personali per una vasta gamma di finalità.

La presente tesi ha esplorato lo sviluppo dei concetti di privacy e sicurezza dei dati personali, presentando alcune delle metodologie di anonimizzazione maggiormente rilevanti, illustrandone applicazioni pratiche oltre ad implicazioni legali.

Le tecniche di anonimizzazione, quali k-anonimity, l-diversity e t-closeness, sono in grado di offrire diversi livelli di protezione dei dati e sono costantemente sottoposte a nuovi studi finalizzati a perfezionarle; Al momento della redazione dell'elaborato, tuttavia, presentano diverse criticità: attacchi di *linkage* mediante analisi incrociata di dataset possono compromettere la sicurezza dei dati e sollevano interrogativi sulla loro efficacia.

Estremamente rilevante e generalmente riconosciuto come un approccio robusto ed efficace è rappresentato dalla *differential privacy*, la quale si distingue da altre tecniche grazie a garanzie matematiche sulla protezione dei dati.
Tuttavia, l'introduzione di concetti come il *rumore* risultano di difficile applicazione: identificare un'adeguata soglia di alterazione delle informazioni, assicurando massimi livelli di utilità del dato e di accuratezza, non è di facile valutazione.

In ambito normativo, GDPR e standard internazionali delineano un quadro di riferimento importante e complesso in merito alla protezione dei dati personali: l'ottemperanza ai principi in essi contenuti richiede un'attenta interpretazione delle norme e una minuziosa analisi del caso concreto. Le recenti pronunce del Garante per la Protezione dei Dati Personali e l'approccio particolarmente stringente adottato dall'autorità evidenziano come non si possa prescindere da una valutazione interdisciplinare delle metodologie da adottare, sottolineando lo stretto legame tra aspetti legali ed ingegneristici.
Regolamenti come l'A.I. Act, inoltre, evidenziano come lo sviluppo del framework legislativo richieda un continuo studio e un costante aggiornamento per quanto concerne la sicurezza dei dati degli individui, al fine di rispondere adeguatamente a rischi emergenti.

Centrale è il delicato trade-off tra utilità dei dati e protezione della privacy degli individui. Da un lato, un alto livello di sicurezza assicura una protezione maggiore della privacy degli individui, limitando però significativamente la capacità di estrarre informazioni rilevanti dai dati. Dall'altro, approcci meno stringenti sono in grado di preservare un maggiore livello di dettaglio, aumentando tuttavia il rischio di re-identificazione e di violazione della privacy. Alla domanda: "Qual è il corretto bilanciamento tra utilità e sicurezza?" non appare quindi esistere una risposta univoca. Ogni misura deve essere adeguatamente contestualizzata e dipende fortemente dalle finalità previste dall'utilizzo delle informazioni, oltre che dalla sensibilità di queste.

In conclusione, l'anonimizzazione dei dati personali richiede un approccio multidisciplinare, combinando competenze tecniche e legali; è fondamentale sviluppare e perfezionare metodologie di anonimizzazione dei dati personali, al contempo promuovendo una "cultura della privacy" secondo un approccio di *privacy-by-design* e sensibilizzando gli utenti stessi in merito alla condivisione dei propri dati; attraverso un impegno congiunto tra legislatore, tecnici e società, è possibile affrontare in modo efficace sfide future in un ambito di tale criticità.

Bibliografia e fonti

Di seguito le fonti utilizzate nel corso dello sviluppo del presente elaborato.

- Samuel D. Warren, Louis D. Brandeis, “The Right To Privacy”, Boston, Dicembre 1890
- The Development of the Theory of the Right to Privacy in France, Wencelas J. Wagner - Indiana University School of Law, 1971
- NIOD Institute for War, Holocaust and Genocide Studies, www.niod.nl
- Trib. Roma, 14 settembre 1953, Caruso c. Soc. p. a. Produzione associata Tirrena Asso film, in Foro it., 77, I, n. 4, 1954, 115 ss.;
- Legge n. 675 del 31 dicembre 1996 “Tutela delle persone e di altri soggetti rispetto al trattamento dei dati personali”
- Regolamento (UE) 2016/679 del Parlamento Europeo e del Consiglio del 27 aprile 2016
- ISO/IEC 27559:2022
- https://www.private-ai.com/en/2023/05/17/iso-iec-27559-2022/
- https://cloud.google.com/sensitive-data-protection/docs/concepts-bucketing?hl=it
- Samarati Pierangela; Sweeney Latanya (1998). "Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression" (PDF). Harvard Data Privacy Lab.
- L. Sweeney, Simple Demographics Often Identify People Uniquely. Carnegie Mellon University, Data Privacy Working Paper 3. Pittsburgh 2000
- Li, Ninghui; Li, Tiancheng; Venkatasubramanian, Suresh (2007). "T-Closeness: Privacy Beyond k-Anonymity and l-Diversity". t-Closeness: Privacy beyond k-anonymity and l-diversity
- Arvind Narayanan, Vitaly Shmatikov, “Robust De-anonymization of Large Sparse Datasets” - The University of Texas at Austin.
- The Algorithmic Foundations of Differential Privacy, da Foundations and Trends R in Theoretical Computer Science Vol. 9, Nos. 3–4 (2014) 211–407 c 2014 C. Dwork and A. Roth DOI: 10.1561/0400000042
- Warner, S. L. (March 1965). "Randomised response: a survey technique for eliminating evasive answer bias". Journal of the American Statistical Association. 60 (309). Taylor & Francis: 63–69. doi:10.1080/01621459.1965.10480775. JSTOR 2283137. PMID 12261830. S2CID 35435339.
- Provvedimento 1° giugno 2023, n. 226, Garante per la Protezione dei Dati Personali
- https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9920562

- Regolamento (UE) 2024/1689 del Parlamento Europeo e del Consiglio del 13 giugno 2024

- James Jordon, Lukasz Szpruch, Florimond Houssiau, Mirko Bottarelli, Giovanni Cherubin, Carsten Maple, Samuel N. Cohen, Adrian Weller - "Synthetic Data - what, why and how?" arXiv:2205.03257 [cs.LG]

- https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

- https://www.gartner.com/en/newsroom/press-releases/2022-06-22-is-synthetic-data-the-future-of-ai

- https://blogs.nvidia.com/blog/american-express-deep-learning/

- https://www.jpmorgan.com/technology/technology-blog/synthetic-data-for-real-insights

- Hradec, J., Craglia, M., Di Leo, M., De Nigris, S., Ostlaender, N. and Nicholson, N., Multipurpose synthetic population for policy applications, EUR 31116 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-53478-5, doi:10.2760/50072, JRC128595.

- Alexander T.P. Boudewijn, Andrea Filippo Ferraris, Daniele Panfilo, Vanessa Cocca, Sabrina Zinutti, Karel De Schepper, Carlo Rossi Chauvenet - "Privacy Measurement in Tabular Synthetic Data: State of the Art and Future Research Directions", arXiv:2311.17453v1 [cs.AI] 29 Nov 2023

- André Bauer, Simon Trapp, Michael Stenger, Robert Leppich, Samuel Kounev, Mark Leznik, Kyle Chard, Ian Foster – "Comprehensive Exploration of Synthetic Data Generation: A Survey", arXiv:2401.02524v2 [cs.LG] 1 Feb 2024

- Theresa Stadler, Bristena Oprisanu, Carmela Troncoso, "Synthetic Data – Anonymisation Groundhog Day", arXiv:2011.07018v6 [cs.LG] 24 Jan 2022

- https://www.priv.gc.ca/en/blog/20221012/ Office of the Privacy Commissioner of Canada

- Maryam Archie, Sophie Gershon, Abigail Katcoff, and Aaron Zeng – "Who's Watching? De-anonymization of Netflix Reviews using Amazon Reviews"

- Vimercati, S.d.C.d., Foresti, S. (2011). Quasi-Identifier. In: van Tilborg, H.C.A., Jajodia, S. (eds) Encyclopedia of Cryptography and Security. Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-5906-5_763

Ringraziamenti

Ogni storia volge al termine, e questa tesi rappresenta la fine di 3 anni in cui la multidisciplinarietà ha decretato il genere di attività che tutt'ora svolgo. La possibilità di spaziare tra diversi ambiti garantisce flessibilità e richiede capacità di adattamento, per cui sono felice di non aver unicamente appreso concetti "scolastici", ma anche un insieme di *soft skills* che, al giorno d'oggi, sono essenziali in qualsiasi tipologia di occupazione. Diritto e Tecnologia è un Corso di Laurea che risponde precisamente a tali necessità.

Tengo a ringraziare la mia Relatrice, la Prof.ssa Claudia Sandei, per avermi trasmesso con grandi professionalità ed umanità la sua stessa passione per tematiche tanto complesse quanto attuali, sin dall'insegnamento "Proprietà Intellettuale, Nuove Tecnologie e Concorrenza", e per avermi seguito con estrema cura e precisione nella stesura del presente elaborato, offrendo preziosi consigli e grande disponibilità.

Ringrazio la mia famiglia, a cui devo tutto, per l'immenso supporto nel corso degli studi e gli stimoli che hanno reso possibile il positivo rendimento durante l'intero percorso. Ogni persona necessita di solide fondamenta per poter costruire qualcosa di grande, e posso indubbiamente ritenermi fortunato a disporne.

Ringrazio i miei compagni di corso, in particolar modo coloro i quali sono sempre stati vicini a me, per aver contribuito in modo essenziale alla creazione di un ambiente piacevole e familiare dove ho avuto il piacere di studiare e frequentare le lezioni. La possibilità di condividere sfide e gioie è un fattore determinante per tracciare ricordi e creare legami che vanno ben oltre le circostanze universitarie.

Grazie a Claudio, DPO e grande conoscitore dell'informatica forense, che con i suoi preziosi consigli e le sue competenze ha reso possibile l'applicazione pratica dei concetti appresi durante gli studi, offrendomi la possibilità di svolgere un tirocinio presso il suo studio e dandomi la possibilità di toccare con mano casi reali, ai quali ho avuto l'onore di poter dare il mio contributo.

Ad un nuovo percorso altrettanto stimolante.