



**UNIVERSITÀ
DEGLI STUDI
DI BERGAMO**

Dipartimento di Ingegneria Gestionale,
dell'Informazione e della Produzione

Corso di laurea in
Ingegneria Informatica

Classe n. L-8

Differential privacy: principio e applicazione

Candidato:
Fabio Filippo
Mandalari

Relatore:
Chiar.mo Prof.
Stefano Paraboschi

Matricola n.
1047426

Anno Accademico
2020/2021

Ai miei genitori

ABSTRACT

I primi passi verso il modello di differential privacy (privacy differenziale) furono compiuti nel 2003, quando Kobbi Nissim e Irit Dinur capirono che la privacy degli individui i cui dati sono memorizzati all'interno di record di un database non può essere protetta senza che venga iniettata una certa quantità di rumore all'interno dei dati stessi. Questa intuizione portò nel 2006 Cynthia Dwork, ricercatrice di Microsoft, alla formalizzazione di una vera e propria teoria riguardante la tecnica di differential privacy. L'assunto di base è che se un database contiene n record e un attaccante conosce i dati contenuti all'interno di $n-1$ record, non deve essere posto nella condizione di poter inferire sui dati contenuti nell'unico record di cui non conosce alcuna informazione. Sotto questa ipotesi, la privacy differenziale garantisce l'esistenza di una bassa probabilità di risalire al dato sconosciuto di interesse.

Al giorno d'oggi la tecnica di differential privacy rappresenta una solida garanzia di privacy individuale, tanto da essere implementata in numerosi ambiti:

- 2008: US Census Bureau, per aver mostrato i modelli di pendolarismo;
- 2014: RAPPORT di Google, per la telemetria come l'apprendimento delle statistiche su software indesiderato che dirotta le impostazioni degli utenti;
- 2015: Google, per la condivisione delle statistiche storiche sul traffico;
- 2016: Apple ha annunciato l'intenzione di utilizzare la privacy differenziale in iOS 10 per migliorare la sua tecnologia di assistente personale intelligente;
- 2017: Microsoft, per la telemetria in Windows;
- 2019: Privtar Lens è un'API che utilizza la privacy differenziale;
- 2020: LinkedIn, per le domande degli inserzionisti.

INDICE

1 Introduzione	6
1.1 La privacy	6
1.2 GDPR, CCPA, LGPD	8
1.2.1 GDPR	8
1.2.2 CCPA	10
1.2.3 LGPD	12
2 Stato dell'arte	15
2.1 Anonimizzazione	15
2.1.1 Anonimizzazione per generalizzazione	17
2.1.1.1 K-anonymity	17
2.1.2 Anonimizzazione per randomizzazione	19
2.1.2.1 Aggiunta di rumore statistico	19
2.1.2.2 Differential privacy	20
2.2 Netflix Prize: case study	22
2.3 Linkage attack	23
2.4 Differential privacy: definizione formale	24
2.4.1 Il parametro ϵ	25
2.4.2 Il parametro δ	26
2.5 Proprietà della differential privacy	27
2.5.1 Privacy budget	27
2.5.1.1 Schema riassuntivo	28
2.5.2 Sensitività	29
2.5.3 Meccanismo di Laplace	30
2.6 GDP, LDP, ESA	31
2.6.1 GDP (Global Differential Privacy)	32
2.6.2 LDP (Local Differential Privacy)	33
2.6.3 ESA (Encode, Shuffle, Analyze)	34

3 Setup environment	35
3.1 Presentazione dei tools	35
3.2 Descrizione del problema	35
3.3 Dalla teoria alla pratica	35
3.3.1 $\epsilon = 0,01$	36
3.3.2 $\epsilon = 1$	39
3.3.3 $\epsilon = 100$	41
3.4 Osservazioni	44
4 Bibliografia	45
5 Sitografia	46

CAPITOLO 1: INTRODUZIONE

1.1 LA PRIVACY

Il termine *privacy* indica il diritto alla riservatezza delle informazioni personali e della propria vita privata. Si tratta di un diritto fondamentale oggi riconosciuto dall'ordinamento giuridico di tutti i paesi europei e delle principali nazioni del mondo.

La sua affermazione come posizione giuridica, tuttavia, ha richiesto un lento processo di riconoscimento. Fino alla fine del 1800 la legge proteggeva esclusivamente il diritto di proprietà e tutelava le persone rispetto alle invasioni fisiche della loro abitazione. Solo alla fine del 1800 è stato riconosciuto il diritto ad essere lasciati soli, ovvero il diritto a impedire ad altre persone di invadere la sfera privata di ognuno di noi.

In un certo senso, la *privacy* è lo strumento attraverso il quale ognuno di noi “può disegnare un confine tra sé stesso e gli altri”. Si tratta di una situazione giuridica che disciplina il modo in cui una persona vive in società nei confronti delle altre. Proprio per questo motivo il concetto stesso di *privacy* e il suo significato nel corso degli anni hanno subito profondi mutamenti, in relazione al mutare della società e, soprattutto, degli strumenti tecnologici utilizzati quotidianamente.

Con l'affermazione delle moderne tecniche di comunicazione e la facilità di diffusione e duplicazione delle informazioni si è compreso che non era più sufficiente proteggere il diritto ad “essere lasciati in pace” e a non subire intromissioni non gradite nella propria vita privata. È diventato sempre più importante evitare che le altre persone potessero abusare delle informazioni riferite ad un soggetto, raccogliendole a sua insaputa e utilizzandole per finalità non consentite dalla legge. Se non venisse garantita questa tutela, ognuno di noi sarebbe sottoposto a pressioni, richieste e potrebbe subire conseguenze negative che limiterebbero fortemente la sua libertà e l'esercizio dei suoi diritti.

Per questo motivo, nel corso del 1900 la privacy ha esteso il suo significato diventando uno strumento giuridico per garantire anche questa specifica situazione. Il punto fermo di questa evoluzione è che ogni persona è titolare del diritto di disporre dei dati che la descrivono e che ne qualificano l'individualità. La privacy è diventata per ognuno il diritto ad esercitare un controllo sulle informazioni che lo riguardano. In questo senso la privacy consiste: a) nel diritto di sapere che qualcun altro sta raccogliendo informazioni sul proprio conto e per quale finalità desidera utilizzarle; b) nel diritto di decidere se si vuole consentire questa raccolta ed utilizzo o se si preferisce negare tale consenso.

Da questa evoluzione del concetto di privacy deriva l'attuale legislazione in materia di dati personali. Per capire il reale significato di queste regole è importante comprendere che la tutela della privacy oggi si occupa principalmente di garantire il diritto fondamentale di esercitare il pieno e consapevole controllo sui propri dati personali. Quando, al giorno d'oggi, si parla di privacy, quindi, non si fa riferimento solo al diritto alla riservatezza, ma anche al diritto di scelta circa l'uso che si vuole che gli altri facciano dei propri dati personali.

1.2 GDPR, CCPA, LGPD

Si dice spesso che il petrolio del XXI secolo non si trovi nel sottosuolo, ma nel web sotto forma di *dato*. Estremizzando la metafora, chi è in grado di estrarre informazioni dai dati così come il petrolio dal sottosuolo ha in mano le chiavi del mondo.

Opportunamente incrociati e aggregati, i dati sono in grado di fornire una miriade di informazioni con un enorme impatto sulla ricerca, lo sviluppo, il business. Si pensi alle politiche in tempo di pandemia: conoscere i dati relativi allo sviluppo del contagio, i dati di accesso agli ospedali, la gestione dei posti in rianimazione e quanto altro concerne questo particolare contesto è di fondamentale importanza per poter gestire efficacemente una emergenza sanitaria. In questo quadro si collocano il GDPR, il CCPA e la LGPD.

1.2.1 GDPR

Il GDPR (*Global Data Protection Regulation*) è un regolamento dell'UE in materia di trattamento e protezione dei dati adottato il 27 aprile 2016, entrato in vigore il 24 maggio dello stesso anno ed operativo a partire dal 25 maggio 2018. Il GDPR detta regole comuni a tutti e 27 i paesi membri dell'UE e si occupa della tutela dei dati personali delle persone fisiche (persone in carne ed ossa), ma non di quelli delle persone giuridiche (imprese, associazioni, pubblica amministrazione, ...).

Per definizione, sono dati personali tutte le informazioni che consentono di identificare una persona fisica direttamente (ad esempio nome e cognome) oppure indirettamente (ad esempio il numero di targa o il codice fiscale), e che forniscono informazioni sulle sue caratteristiche fisiche, il suo stile di vita, le sue relazioni, il suo stato di salute, la sua situazione economica e sociale, la sua religione o l'adesione ad un partito politico.

Non sono dati personali né i dati anonimi né i dati anonimizzati, ovvero quei dati resi anonimi in un momento successivo rispetto alla loro pubblicazione. Queste due categorie di dati, per come sono state concepite, escludono a priori che il loro uso permetta di risalire alle identità delle persone da cui provengono, dunque non sono soggette al GDPR. Oltre alle categorie dei dati anonimi e di quelli anonimizzati è possibile individuare quella dei dati pseudonimizzati. I dati pseudonimizzati, se presi da soli, non consentono di risalire all'identità della persona che li ha generati, ma se vengono uniti ad altre informazioni, allora sì. In virtù di questa reversibilità dell'anonimato, i dati pseudonimizzati sono trattati al pari dei dati personali, e dunque sono soggetti al GDPR.

Ai sensi del GDPR, i dati personali possono essere divisi in due grandi gruppi:

- Dati personali comuni. Questo tipo di dati sono detti “comuni” perché non rivelano aspetti intimi della vita delle persone. Ne sono un esempio i dati anagrafici, il codice fiscale, il numero di targa e l’indirizzo IP.
- Categorie particolari di dati personali. Questo tipo di dati rivelano aspetti particolarmente intimi e personali della vita delle persone. Appartengono a questa categoria l’origine razziale o etnica, lo stato di salute, la vita sessuale, le opinioni politiche, eventuali trascorsi giudiziari, l’appartenenza ad un sindacato, la confessione religiosa, i dati genetici e quelli biometrici. Questo gruppo in origine conteneva tutti quei dati cosiddetti sensibili.

È fondamentale conoscere questa suddivisione dei dati personali perché il GDPR pone come regola generale il divieto di utilizzo delle categorie particolari di dati personali. L’uso di questo tipo di dati è da considerarsi del tutto eccezionale e deve obbedire a standard di sicurezza più elevati, come può esserlo la richiesta di un consenso esplicito dell’interessato per un ben definito trattamento.

Il GDPR fornisce una definizione specifica per quanto riguarda il concetto di trattamento dei dati. Costituisce un trattamento qualsiasi operazione applicata ai dati personali, compiuta con o senza l’aiuto di processi automatizzati. Esempi di trattamento di dati, quindi, possono essere la raccolta, la suddivisione e la memorizzazione degli stessi.

Per il trattamento dei dati personali il GDPR detta varie regole:

- Il trattamento deve essere lecito.
- Bisogna seguire il cosiddetto principio di minimizzazione del trattamento, secondo il quale bisogna raccogliere il minor numero di dati possibile. In linea di principio, bisogna raccogliere solo i dati strettamente necessari per il particolare scopo.
- I dati vanno conservati per un periodo di tempo limitato. Tale tempo è pari al tempo necessario al raggiungimento del particolare scopo per cui sono stati raccolti. Volendoli conservare per un periodo più lungo, occorre anonimizzarli.
- Il trattamento deve essere trasparente. Il principale mezzo di trasparenza è l’informativa, ovvero quel documento che contiene la spiegazione di tutte le modalità e gli scopi con cui verranno successivamente trattati i dati.
- La conservazione dei dati deve essere sicura, ovvero tale da evitare la perdita, la distruzione o la diffusione indebita dei dati personali.

1.2.2 CCPA

Il CCPA (*California Consumer Privacy Act*) è una legge statale sulla protezione dei dati che disciplina il modo in cui le aziende di tutto il mondo sono autorizzate a trattare le informazioni personali dei cittadini residenti in California. La data della sua entrata in vigore è il 1° gennaio 2020 e si tratta della prima legge di questo genere negli Stati Uniti. Ai sensi del CCPA:

- Per consumatore si intende una persona fisica residente in California;
- Per dati personali si intendono informazioni che identificano, che si riferiscono, che descrivono, che possono essere associate o che potrebbero ragionevolmente essere collegate, direttamente o indirettamente, a uno specifico consumatore. Sono dati personali il nome, l'indirizzo postale, l'indirizzo IP, l'indirizzo e-mail, numero di patente di guida, numero di passaporto, dati biometrici, dati di geolocalizzazione, informazioni professionali o relative all'impiego, cronologia di navigazione, ...
- Per azienda si intende un'organizzazione a scopo di lucro che raccoglie le informazioni personali dei consumatori, ne determina le finalità e il metodo di trattamento, si rivolge ai residenti californiani (indipendentemente dal fatto che l'azienda abbia o meno sede in California) e soddisfi almeno uno dei requisiti:
 - Ha un fatturato annuo lordo superiore a 25 milioni di dollari;
 - Almeno il 50% del proprio fatturato deriva dalla vendita di dati personali;
 - Acquista, riceve, vende o condivide ogni anno per finalità commerciali le informazioni personali di 50.000 o più consumatori.

Il CCPA viene applicato quando una particolare attività commerciale si rivolge ad utenti californiani e prevede i seguenti diritti per i consumatori:

- Diritto ad essere informati. I consumatori hanno il diritto di essere informati sulle modalità di trattamento dei loro dati prima o durante l'atto di raccolta degli stessi.
- Diritto di accesso. I consumatori che fanno una richiesta verificabile dall'azienda stessa hanno il diritto di accedere ai propri dati personali.
- Diritto alla portabilità dei dati. Nel caso in cui un'azienda esaudisca la richiesta di accesso in formato elettronico, le informazioni devono essere fornite al consumatore in un formato facilmente fruibile, portabile e che consenta al consumatore di trasmettere queste informazioni ad un'altra azienda senza difficoltà.
- Diritto alla cancellazione. Il CCPA garantisce ai consumatori il diritto di richiedere la cancellazione di qualsiasi dato personale raccolto sul loro conto. Una volta ricevuta una richiesta (verificabile) di cancellazione, l'attività commerciale deve cancellare i dati personali del consumatore richiedente dai propri registri e chiedere a tutti i fornitori di servizi correlati di fare altrettanto.
- Diritto di opporsi. Ai sensi del CCPA il consumatore ha il diritto di impedire in ogni momento e con una semplice comunicazione all'azienda la vendita dei propri dati personali a terzi. Oltre alla vendita sono contemplate le casistiche di cessione, rilascio, divulgazione, diffusione, comunicazione orale, per iscritto o mediante mezzi elettronici.
- Diritto all'opt-in. Le aziende non possono vendere i dati personali di un consumatore di età inferiore a 16 anni a meno che:
 - Il consumatore abbia tra i 13 e i 16 anni e abbia effettuato l'opt-in;
 - Il consumatore ha meno di 13 anni, ma un suo genitore/tutore ha effettuato l'opt-in in sua vece.
- Diritto a non essere discriminati. Il CCPA vieta alle aziende di discriminare i consumatori per aver esercitato i diritti conferiti loro dalla legge. In particolare, per una azienda non è possibile:
 - Negare beni o servizi ai consumatori;
 - Applicare prezzi o tariffe diverse per beni o servizi;
 - Fornire beni o servizi con diverso livello di qualità;
 - Suggestire che il consumatore riceverà beni o servizi ad un prezzo diverso.

1.2.3 LGPD

La LGPD (*Lei Geral de Proteção de Dados*) è la legge brasiliana sulla protezione dei dati personali. In un certo senso può essere pensata come la risposta del Brasile al GDPR. La LGPD è pienamente applicabile dal 18 settembre 2020 e intende sostituire o completare l'attualmente frammentato panorama giuridico composto da più di 40 norme federali con un quadro normativo principale. In termini pratici, si applica quando:

- Le attività di trattamento dati sono svolte in Brasile (ad esempio, si usano dei server con sede in Brasile);
- Vengono offerti beni o servizi a persone situate in Brasile, indipendentemente dalla loro nazionalità;
- Vengono trattati dati di persone che si trovano in Brasile (anche se queste si trovavano in Brasile solo al momento della raccolta dei dati, e poi si sono spostate).

La definizione che la LGPD dà di dato personale è piuttosto ampia. Appartengono alla categoria dei dati personali tutte le informazioni che sono riconducibili a un individuo identificato o identificabile. Al pari del GDPR, i dati pseudonimizzati sono da considerarsi a tutti gli effetti dati personali, mentre i dati anonimi e quelli anonimizzati no. Esempi di dati personali sono i nomi delle persone, i dati biometrici o quelli inerenti alla salute, dati relativi alla navigazione web e gli indirizzi IP, e-mail personali, dati relativi alle opinioni politiche e dati sull'orientamento sessuale. Vi è inoltre la medesima suddivisione tra dato personale comune e dato personale sensibile. I dati sensibili, potendo esporre più facilmente l'utente a rischi di discriminazione, devono essere trattati con ulteriori livelli di sicurezza e con basi giuridiche molto specifiche. In generale, si possono trattare dati sensibili solo se l'utente ha dato il proprio consenso per il particolare trattamento per cui sono richiesti.

I principi per il trattamento dei dati personali sono molto simili a quelli del GDPR:

- Il trattamento deve avere una finalità. Ciò significa che qualsiasi attività di trattamento dei dati deve essere svolta per scopi legittimi, specifici, espliciti e chiaramente comunicati. Non è possibile effettuare alcun trattamento aggiuntivo che non sia in linea con le finalità originarie comunicate.
- Adeguatezza. Sia il modo di tracciare i dati sia i dati trattati devono essere in linea con le finalità del trattamento.
- Limitazione delle finalità. Simile al concetto di minimizzazione dei dati ai sensi del GDPR, significa che bisogna limitarsi a trattare solo i dati necessari per il raggiungimento delle proprie finalità.
- Libertà di esercitare i propri diritti e di accedere ai propri dati. gli utenti devono poter esercitare liberamente i loro diritti ai sensi della LGPD e devono poter accedere liberamente e gratuitamente a tutte le informazioni relative al trattamento dei propri dati personali.
- Trasparenza. Le informazioni sul trattamento dei dati devono essere chiare, precise e facilmente accessibili agli utenti.
- Sicurezza. Sia il titolare del trattamento sia gli eventuali responsabili del trattamento devono essere sicuri di disporre di misure tecniche e organizzative che proteggono i dati personali da accessi non autorizzati, distruzione accidentale o illecita, perdita, alterazione e comunicazione o diffusione non autorizzata.
- Prevenzione. È responsabilità del titolare e del responsabile del trattamento adottare misure tecniche e organizzative atte a prevenire eventuali danni causati dal trattamento dei dati personali.
- Non discriminazione. Nessun trattamento dei dati deve avvenire con finalità discriminatorie.
- Responsabilità. In qualità di titolare del trattamento dei dati, si è tenuti a rispettare la legge e bisogna essere in grado di dimostrarlo.

Ai sensi della LGPD, gli utenti hanno il diritto di:

- Conferma. Gli utenti hanno il diritto di ricevere conferma dell'esistenza di un trattamento.
- Accesso. Gli utenti hanno il diritto di accedere ai propri dati trattati dal titolare.
- Portabilità dei dati. Su espressa richiesta, gli utenti hanno il diritto di trasferire i propri dati ad un altro fornitore di servizi o prodotti, in conformità alle norme dell'autorità nazionale e nel rispetto del segreto commerciale e industriale.
- Rettifica. Gli utenti hanno il diritto di richiedere la rettifica dei propri dati personali se sono imprecisi o incompleti.
- Anonimizzazione. Gli utenti hanno il diritto all'anonimizzazione, al blocco o all'eliminazione dei dati personali non necessari o in eccesso, o di qualsiasi dato non trattato in conformità alla LGPD.
- Cancellazione/oblio. Gli utenti hanno il diritto di far cancellare i loro dati personali se il trattamento di tali dati è basato sul consenso.
- Essere informati. Gli utenti hanno il diritto di essere informati sui responsabili del trattamento e sulle terze parti che accedono o trattano i suoi dati personali.
- Revoca. Gli utenti hanno il diritto di revocare o ritirare il proprio consenso.
- Presentare reclami. Gli utenti hanno il diritto di presentare un reclamo all'Autorità per la protezione dei dati.
- Opporsi. Gli utenti hanno il diritto di opporsi al trattamento dei propri dati personali in caso di mancato rispetto delle disposizioni di legge.
- Richiesta di revisione. Gli utenti hanno il diritto di richiedere la revisione delle decisioni prese sulla base di un trattamento automatizzato dei dati riguardanti i loro interessi.

CAPITOLO 2: STATO DELL'ARTE

2.1 ANONIMIZZAZIONE

Il processo di anonimizzazione costituisce un trattamento successivo dei dati personali rispetto a quello effettuato per la finalità originaria perseguita dal titolare e si pone come obiettivo quello di pervenire ad una nuova rappresentazione del dato anonimizzato.

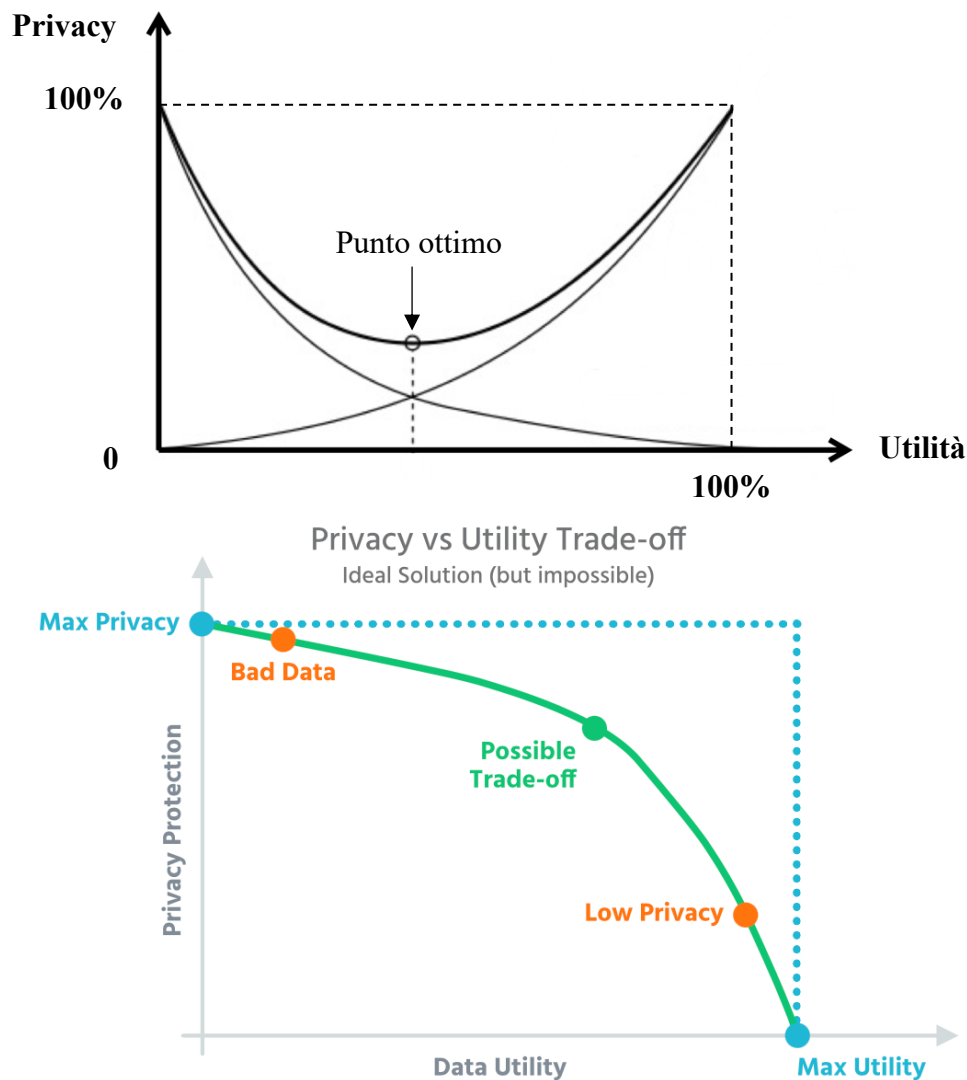
Quando si è parlato del concetto di privacy si è detto: “se non venisse garantita questa tutela, ognuno di noi sarebbe sottoposto a pressioni, richieste e potrebbe subire conseguenze negative che limiterebbero fortemente la sua libertà e l’esercizio dei suoi diritti”. Ciò detto, quando si scongiura ogni impatto sulla persona? Un efficace processo di anonimizzazione scongiura impatti sulla persona solo ed esclusivamente se è in grado di impedire a chiunque disponga di un insieme di dati anonimizzati di: 1) isolare una persona in un gruppo, 2) collegare un dato anonimizzato a dati riferibili ad una persona presenti in un distinto insieme di dati, 3) dedurre da un dato anonimizzato nuove informazioni riferibili ad una persona.

Su un piano più operativo, questi obiettivi possono essere perseguiti mediante l’applicazione, eventualmente congiunta, di diverse tecniche di anonimizzazione che, in generale, sono raggruppabili in due categorie: la generalizzazione e la randomizzazione dei dati, entrambe concepite per introdurre un grado di incertezza misurabile in termini probabilistici sull’attribuzione di un dato anonimizzato a un determinato soggetto.

La *generalizzazione* consiste nel diluire il valore assunto dagli attributi dei dati modificandone la scala o l’ordine di grandezza (ad esempio, una regione anziché una città, un mese anziché una settimana). L’incertezza, in questo caso, è legata al fatto che quanto più lasca è la scala dei valori degli attributi, tanto maggiore è il numero di interessati potenzialmente riferibili ad un certo attributo “generalizzato”.

La *randomizzazione* è una famiglia di tecniche che modifica la veridicità dei dati al fine di eliminare, ove possibile, il legame che esiste tra il dato puntuale e la persona. Se, infatti, i dati vengono resi sufficientemente incerti, essi non possono più essere riferiti ad una persona specifica, a tal punto da trasferire in taluni casi questa incertezza persino alla stessa presenza di un dato riferibile ad un interessato all’interno di un database.

Naturalmente, per l'impiego di entrambe le tecniche, si pone un problema di utilità del dato a seguito del processo di anonimizzazione. Nel caso delle tecniche che operano per generalizzazione, se la scala è troppo lasca, il dato rischia di perdere ogni valenza semantica, diventando inidoneo a esprimere qualsiasi nesso di correlazione utile a descrivere un fenomeno. Nel caso delle tecniche che operano per randomizzazione, se il rumore prevale rispetto al dato utile, questo diventa, oltre che incerto (ossia non riferibile ad alcuno), inaccurato e inidoneo a qualsiasi tipo di analisi. È dunque necessaria una calibrazione della distorsione in ragione dell'uso che si vorrà fare del dato anonimizzato. Nell'ambito degli studi sui sistemi di privacy preserving si definisce il problema del *privacy vs utility trade-off* come quel problema tale per cui tanto maggiore è l'utilità che si vuole trarre da un dato, quanto minore deve essere la privacy preservata:



In questo contesto, dunque, il termine “utilità” è da leggersi in termini di “qualità”.

2.1.1 ANONIMIZZAZIONE PER GENERALIZZAZIONE

La famiglia delle tecniche di anonimizzazione che opera per generalizzazione ha come obiettivo quello di rendere meno dettagliati gli attributi delle persone interessate presenti in una tabella, modificando la rispettiva scala o l'ordine di grandezza in modo che più righe di quella tabella presentino la stessa combinazione di attributi generalizzati.

2.1.1.1 K-ANONYMITY

K-anonymity è una tecnica che rientra nella famiglia delle tecniche di anonimizzazione che operano per generalizzazione la cui idea è quella di rendere meno precisi e dettagliati i record all'interno di un database. Nel k-anonimato, la "k" indica un numero che rappresenta le dimensioni di un gruppo. Se per qualsiasi individuo presente nel database ci sono almeno k-1 individui con le stesse proprietà, significa che quel database ha raggiunto il k-anonimato. Con l'obiettivo di generalizzare le singolarità, quindi, la tecnica k-anonimato si prefigge di scongiurare l'individuazione di specifiche persone mediante un loro raggruppamento con almeno altre k persone.

Suppongo di disporre del seguente database non anonimizzato:

Nome	Sesso	Età	Domicilio	Religione	Malattia
John	M	29	Londra	Cristiana	No malattia
Marc	M	24	Parigi	Musulmana	Diabete
Joe	M	28	Parigi	Cristiana	Infezione virale
George	M	27	Roma	Cristiana	Diabete
Jack	M	24	Londra	Cristiana	Infezione virale
Will	M	23	Roma	Indu	Ipertensione
Oscar	F	19	Roma	Musulmana	Ipertensione
Emily	M	29	Londra	Cristiana	Diabete
Amelia	F	17	Roma	Indu	No malattia
Harry	F	19	Parigi	Musulmana	Diabete
Thom	F	20	Parigi	Indu	Infezione virale

Volendo proteggere l'informazione sensibile dello stato di salute dei soggetti presenti preservando allo stesso tempo la significatività dei record, la tecnica k-anonymity prevede che si seguano due passi:

- Sostituire con il simbolo “*” i valori di quegli attributi che veicolano informazioni secondarie per il particolare scopo (in questo caso la protezione dell'informazione sullo stato di salute dei soggetti). Considero gli attributi Nome e Religione.
- Generalizzare i valori degli attributi che possono avere una certa valenza nei confronti dell'informazione che si vuole proteggere sostituendo il valore effettivo con un intervallo più o meno ampio. Considero l'attributo Età.

In questo caso è ottenuta la proprietà 2-anonymity rispetto agli attributi Sesso, Età e Domicilio: per ogni terna di valori degli attributi Sesso-Età-Domicilio esistono almeno due record che si differenziano per il valore dell'attributo Malattia:

Nome	Sesso	Età	Domicilio	Religione	Malattia
*	M	$20 \leq \text{età} < 30$	Londra	*	No malattia
*	M	$20 \leq \text{età} < 30$	Parigi	*	Diabete
*	M	$20 \leq \text{età} < 30$	Parigi	*	Infezione virale
*	M	$20 \leq \text{età} < 30$	Roma	*	Diabete
*	M	$20 \leq \text{età} < 30$	Londra	*	Infezione virale
*	M	$20 \leq \text{età} < 30$	Roma	*	Iipertensione
*	F	$\text{età} \leq 20$	Roma	*	Iipertensione
*	M	$20 \leq \text{età} < 30$	Londra	*	Diabete
*	F	$\text{età} \leq 20$	Roma	*	No malattia
*	F	$\text{età} \leq 20$	Parigi	*	Diabete
*	F	$\text{età} \leq 20$	Parigi	*	Infezione virale

2.1.2 ANONIMIZZAZIONE PER RANDOMIZZAZIONE

La famiglia delle tecniche di anonimizzazione che opera per randomizzazione si compone di un insieme di tecniche che hanno l'obiettivo di attenuare il legame esistente tra la persona ed i dati che la riguardano, in modo che, noti i dati randomizzati, non si possa risalire alla persona cui si riferiscono. Il termine randomizzazione indica che questa operazione è compiuta introducendo un qualche elemento di casualità nei dati.

Le tecniche più importanti di questa famiglia sono due:

- Aggiunta di rumore statistico;
- Differential privacy.

Nella pratica esistono anche altre tecniche, come ad esempio la permutazione e i questionari polarizzati, ma la loro trattazione non è contemplata in questa opera.

2.1.2.1 AGGIUNTA DI RUMORE STATISTICO

Per rumore statistico si intende una variabile aleatoria che, se sommata al valore reale del dato, perturba l'informazione estraibile rendendola meno accurata. Il rumore viene generato mediante un programma automatico, detto generatore di numeri pseudocasuali. Il livello di rumore determina, da un lato, il livello di mascheramento dell'informazione vera (quanto più rumore viene aggiunto, tanto più l'informazione singola viene protetta), ma d'altro canto diminuisce l'accuratezza dell'informazione estraibile. È evidente, dunque, che l'aggiunta di rumore debba essere realizzata nel rispetto di alcuni vincoli:

- La distribuzione complessiva dei valori deve rimanere inalterata. Bisogna cioè fare in modo che i valori aggregati risultino statisticamente uguali se calcolati con i valori rumorosi piuttosto che con quelli reali.
- La gamma di valori dell'attributo ed eventuali vincoli logici derivanti da altri attributi devono essere rispettati. Per esempio, nel caso delle età il rumore aggiunto non deve essere tale da fornire età negative oppure troppo elevate.

Ciò detto, questa tecnica è vulnerabile. Se il rumore viene aggiunto in maniera nuova ad ogni interrogazione del database, interrogando molte volte il database si ottengono molte risposte diverse, che però hanno lo stesso valore medio (cioè il valore vero). Calcolando la media delle risposte fornite si ottiene una stima del valore vero. Stando a questo ragionamento, maggiore è il numero di interrogazioni (e quindi il numero di risposte rumorose), maggiore è l'accuratezza della stima.

2.1.2.2 DIFFERENTIAL PRIVACY

In molti casi la violazione della privacy può avvenire anche se l'accesso al database si limita ad interrogazioni aggregate. Questo è il caso delle cosiddette basi di dati statistiche, ovvero delle banche dati dalle quali è permessa solamente l'estrazione di informazioni statistiche aggregate mediante interrogazioni riguardanti non uno specifico record, ma un insieme di record. Nonostante le limitazioni imposte da questo tipo di database è comunque possibile in certi casi risalire al singolo individuo.

Suppongo, ad esempio, di poter interrogare un database statistico sportivo che, per semplicità, contiene solo il nome delle persone e un indicatore che segnala se la specifica persona ha mai giocato a calcio in vita sua. Il database in questione è del tipo:

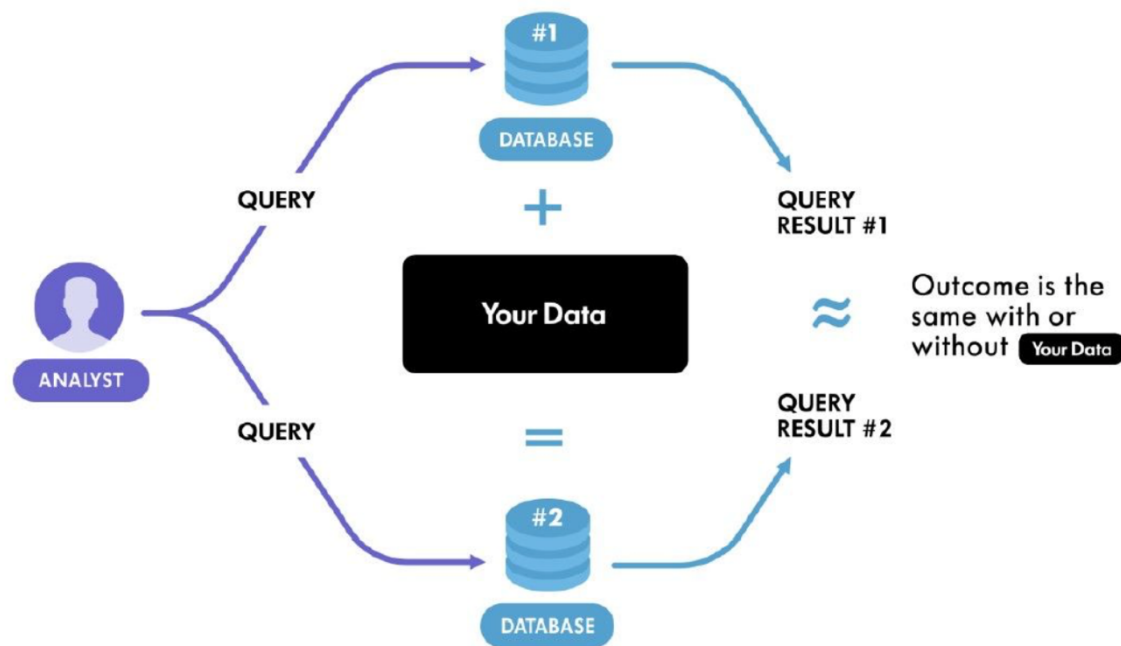
Nome	Calcio?
Andrea	SÌ
Beatrice	NO
Carlo	NO
Daniele	SÌ
Martina	NO

Supponendo che a questo database si possano sottoporre solamente interrogazioni del tipo “quante persone tra le prime n hanno giocato a calcio almeno una volta nella loro vita?”, mi concentro sul voler risalire a Daniele. Non è ovviamente possibile effettuare un'interrogazione specifica perché la risposta violerebbe la privacy di Daniele e ciò non sarebbe permesso. Tuttavia, sapendo che l'anagrafica di Daniele lo colloca in quarta posizione, basta mandare in esecuzione in successione le due seguenti interrogazioni aggregate “quante persone tra le prime tre hanno giocato a calcio almeno una volta in vita loro?” e “quante persone tra le prime quattro hanno giocato a calcio almeno una volta in vita loro?”, e poi calcolare la differenza tra i due risultati. Se questa differenza vale 0, allora la persona in quarta posizione, ovvero Daniele, non ha mai giocato a calcio. Viceversa, se il risultato è 1, la persona in quarta posizione ha giocato a calcio almeno una volta in vita sua. Nel caso in questione, la prima interrogazione fornisce come risultato 1, mentre la seconda 2. La differenza tra i due valori è 1, e quindi si deduce che Daniele ha giocato a calcio almeno una volta in vita sua.

Occorre quindi individuare un meccanismo di risposta ad interrogazioni di questo tipo che non permetta tali deduzioni.

Per risolvere questo problema è stata introdotta la tecnica di *differential privacy*.

Un algoritmo garantisce un certo livello di privacy differenziale se, per due qualsiasi database che differiscono per un solo record, le risposte fornite dall'algoritmo alla stessa interrogazione sottoposta ad entrambi i database sono sostanzialmente le stesse. Per “sostanzialmente le stesse” si intende che la probabilità che siano diverse è molto bassa. Se le risposte sono indistinguibili, certamente non possono essere sfruttate per effettuare deduzioni sull'unico record che è presente in uno solo dei due database.



2.2 NETFLIX PRIZE: CASE STUDY

Dal momento del suo lancio nel lontano 1997, il core business di Netflix è consistito nel noleggio di DVD e videogiochi prenotabili online e recapitati direttamente a casa. Già allora era stato implementato un sistema di algoritmi per inviare agli utenti consigli su nuovi titoli basati sullo storico delle loro preferenze. Fu però con il lancio della piattaforma per la trasmissione di film in streaming che tali algoritmi divennero la chiave per le strategie aziendali. Infatti, mentre in precedenza gli unici dati a disposizione erano quelli relativi ai DVD noleggiati, adesso si ha accesso a molte più informazioni estremamente dettagliate: i film visti, le serie TV iniziate, le serie TV iniziate e non finite, le serie TV divorate, le serie TV iniziate e riprese dopo un certo tempo, ...

Nel 2006 l'azienda statunitense indisse un concorso con un premio di un milione di dollari per chi fosse riuscito a migliorare l'accuratezza di Cinematch, l'algoritmo per il suggerimento dei film utilizzato all'epoca, di almeno 10 punti percentuali.

Tra le proposte più innovative ci fu quella del team BellKor's Pragmatic Chaos, che nel 2009 si aggiudicò il premio. Il team elaborò 107 algoritmi in oltre 2000 ore di lavoro.

Per questa occasione Netflix pubblicò un database anonimizzato contenente 100 milioni di recensioni di film create da 500.000 utenti, l'ID associato a ciascun utente e la data in cui ogni utente ha rilasciato la recensione. Il dataset fu anonimizzato per garantire la privacy degli utenti che avevano rilasciato la recensione.

Narayanan e Shmatikov, due ricercatori dell'università del Texas ad Austin, riuscirono a risalire ai dati reali degli utenti. I due effettuarono uno scraping dei dati presenti su IMDB (Internet Movie DataBase) estraendo le recensioni dei film. Ciò che fecero in pratica è estrarre da IMDB, un sito web di proprietà di Amazon che gestisce informazioni su film, attori, registi, recensioni, sunto delle trame, ..., i dati degli utenti che avevano rilasciato almeno una recensione e la recensione stessa. Successivamente, avvalendosi di avanzate tecniche statistiche e informatiche, incrociarono i dati (pubblici) estratti mediante scraping con quelli anonimi rilasciati da Netflix. Questa operazione permise di de-anonimizzare una grande porzione di dati.

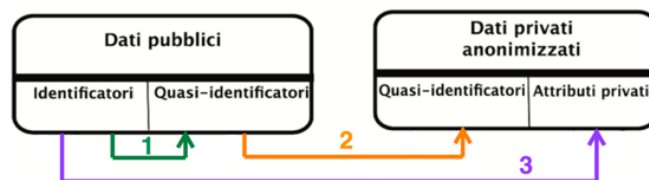
Il punto della situazione è chiaro: una tecnica apparentemente ricercata come l'anonimizzazione dei dati è semplicemente insufficiente per garantire la privacy dei soggetti i cui dati sono memorizzati in un database.

2.3 LINKAGE ATTACK

Un attacco linkato è un attacco informatico alla privacy delle persone che ha come scopo la re-identificazione degli individui presenti all'interno di un database contenente dati anonimizzati. L'attacco viene effettuato combinando mediante join le informazioni presenti all'interno del database con delle altre provenienti da altre fonti.

Per procedere con la trattazione bisogna introdurre il concetto di quasi-identificatore. Per quasi-identificatori si intende qualsiasi combinazione di caratteristiche della persona utile a identificarla. Alcuni quasi-identificatori, come ad esempio il nome, il cognome, la data di nascita e il codice fiscale rendono la persona direttamente identificabile. In tal caso, non si parla di quasi-identificatori, bensì di identificatori. Nel contesto delle basi di dati, il concetto di quasi-identificatore viene a coincidere con quello di attributo.

I macro-passi da seguire per effettuare un attacco linkato sono:



In una fase di inizializzazione della procedura di attacco, l'attaccante, avendo un target di riferimento, individua tutti i possibili identificatori disponibili pubblicamente che lo possono aiutare a portare a termine l'attacco. Con questi crea un database ad-hoc (fase 0). Disponendo del database statistico, l'attaccante sa quali sono i quasi-identificatori appartenenti al gruppo degli identificatori individuati durante la fase 0 da linkare per sferrare l'attacco (fase 1). In altre parole, i quasi-identificatori individuati dall'attaccante durante la fase 1 sono tutti quegli attributi che l'attaccante sa per certo siano presenti all'interno del database statistico e che, a seguito di ulteriori lavorazioni, lo porterebbero a capire qual è il record del database che descrive il target nella sua totalità qualora riuscisse a portare a termine l'attacco.

Essendo presenti sia il database statistico sia la base di dati creata ad-hoc, l'attaccante può metterle in relazione sfruttando la clausola join con un predicato opportunamente settato basato sui valori assunti dai quasi-identificatori (fase 2).

L'effetto ottenuto dall'esecuzione della join è quello di mettere in diretto collegamento gli identificatori selezionati dai dati pubblici con gli attributi privati presenti all'interno del database statistico, cosicché si venga a creare una corrispondenza univoca tra ciascun dato identificatore pubblico (visibile) e il rispettivo dato privato (anonimo) (fase 3).

2.4 DIFFERENTIAL PRIVACY: DEFINIZIONE FORMALE

Prima di procedere fornendo la definizione operativa di privacy differenziale è necessario introdurre tutti gli elementi che rientrano nella definizione stessa dandogli un significato:

- $K \Rightarrow$ Algoritmo randomizzato, ovvero l'algoritmo posto a supervisione del database che ha il compito di valutare la specificità delle query per capire quanto rumorose devono essere le risposte;
- D_1 e $D_2 \Rightarrow$ Due qualunque database che differiscono al più per un record;
- $K(D_1)$ e $K(D_2) \Rightarrow$ Risposte rumorose fornite dall'algoritmo K a seguito della valutazione della query sia da parte di D_1 sia da parte di D_2 ;
- $S \Rightarrow$ Insieme di tutte le possibili risposte lecite e rumorose che possono essere fornite dall'algoritmo K a seguito della valutazione della query sia da parte di D_1 sia da parte di D_2 .

A questo punto si può procedere fornendo la definizione di ϵ -differential privacy.

Un algoritmo randomizzato K fornisce un livello ϵ di privacy differenziale se:

$$\mathbb{P}[K(D_1) \in S] \leq e^\epsilon \cdot \mathbb{P}[K(D_2) \in S]$$

La chiave di lettura della definizione è: disponendo di due database simili D_1 e D_2 , la probabilità che l'algoritmo K applicato sia a D_1 sia a D_2 dia un certo risultato $K(D_1)$ o $K(D_2)$ tra tutti i possibili risultati S , è più o meno la stessa.

Dal concetto di ϵ -differential privacy prende piede quello di (ϵ, δ) -differential privacy.

Un algoritmo randomizzato K fornisce un livello (ϵ, δ) di privacy differenziale se:

$$\mathbb{P}[K(D_1) \in S] \leq e^\epsilon \cdot \mathbb{P}[K(D_2) \in S] + \delta$$

La prima versione, ovvero quella per cui $\delta = 0$, è anche detta privacy differenziale *pura*, mentre la seconda, ovvero quella per cui $\delta \neq 0$, privacy differenziale *approssimata*.

I parametri ϵ e δ meritano una trattazione approfondita.

2.4.1 IL PARAMETRO ϵ

Per semplicità richiamo la formula per il calcolo delle probabilità secondo l'approccio della privacy differenziale pura:

$$\mathbb{P}[K(D_1) \in S] \leq e^\epsilon \cdot \mathbb{P}[K(D_2) \in S], \text{ da cui: } \frac{\mathbb{P}[K(D_1) \in S]}{\mathbb{P}[K(D_2) \in S]} \leq e^\epsilon$$

Il parametro ϵ è una sorta di manopola che consente di spostare l'attenzione o verso la privacy delle persone i cui dati sono immagazzinati nel database o verso l'accuratezza dell'analisi dei dati stessi.

In particolare, se:

- $\epsilon \rightarrow 0 \Rightarrow e^\epsilon \rightarrow 1 \Rightarrow \mathbb{P}[K(D_1) \in S] \approx \mathbb{P}[K(D_2) \in S] \Rightarrow$ Le probabilità che le risposte anonimizzate fornite dall'algoritmo K a seguito di una interrogazione sia di D_1 sia di D_2 per mezzo di una stessa query, pur essendo diverse, sono così tanto simili da poter essere considerate indistinguibili. Se le risposte fornite da due database che differiscono per un solo record sono indistinguibili, vuol dire che la privacy dell'individuo presente in uno solo dei due database è totalmente preservata. Richiamando il concetto di trade-off che sussiste tra privacy e utilità del dato a fini statistici, in questo caso la privacy dell'individuo presente in uno solo dei due database è totalmente preservata, ma a livello di analisi dei dati si sta sacrificando troppa informazione.
- $\epsilon \rightarrow \infty \Rightarrow e^\epsilon \rightarrow \infty \Rightarrow \mathbb{P}[K(D_1) \in S] \leq e^\epsilon \cdot \mathbb{P}[K(D_2) \in S] \Rightarrow$ Le probabilità che le risposte anonimizzate fornite dall'algoritmo K a seguito di una interrogazione sia di D_1 sia di D_2 per mezzo di una stessa query sono estremamente diverse. In particolare, differiscono per un fattore pari a e^ϵ . Se le risposte fornite da due database che differiscono per un solo record sono estremamente diverse a causa del fatto che uno dei due database contiene un record in più, vuol dire che il fattore di incidenza derivante dalla presenza di quel record è alto e ha un forte impatto nel computo della probabilità. Richiamando il concetto di trade-off che sussiste tra privacy e utilità del dato a fini statistici, in questo caso si ha un sacrificio della privacy dell'individuo presente in uno solo dei due database in favore di una più accurata analisi dei dati.

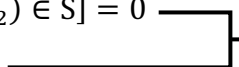
Non esiste una regola scritta per la scelta del valore di ϵ . Valori noti in letteratura sono $\ln(2)$, $\ln(3)$, $1/10$, $1/100$.

2.4.2 IL PARAMETRO δ

Per semplicità richiamo la formula per il calcolo delle probabilità secondo l'approccio della privacy differenziale approssimata:

$$\mathbb{P}[K(D_1) \in S] \leq e^\epsilon \cdot \mathbb{P}[K(D_2) \in S] + \delta$$

Per contestualizzare il significato attribuito al parametro δ assumo:

- $\mathbb{P}[K(D_2) \in S] = 0$
 - $\delta \neq 0$
- 
- $\mathbb{P}[K(D_1) \in S] \leq \delta$

In questa situazione si può notare che, nonostante $\mathbb{P}[K(D_2) \in S] = 0$, sottoponendo a D_1 la stessa query che porta ad avere $\mathbb{P}[K(D_2) \in S] = 0$, si ottiene $\mathbb{P}[K(D_1) \in S] = \delta \neq 0$. In altre parole, secondo questa definizione di privacy differenziale esiste sempre una probabilità δ tale per cui è consentito dire che, pur ottenendo una risposta anonimizzata dall'algoritmo K ad una query, si possa risalire al dato reale partendo da quello anonimo. La differenza sostanziale tra la definizione di ϵ -differential privacy e quella di (ϵ, δ) -differential privacy sta nel fatto che, mentre con la prima e con la medesima assunzione sopra riportata $\mathbb{P}[K(D_2) \in S] = 0$ si sarebbe ottenuto $\mathbb{P}[K(D_1) \in S] = 0$, la seconda, includendo il parametro δ , fa sì che nel computo della probabilità $\mathbb{P}[K(D_1) \in S]$ si debba tenere conto anche dell'eventuale presenza di agenti esterni rispetto alla query che viene sottoposta al database che causano una fuga inaspettata di informazione.

Il parametro δ , dunque, rappresenta l'eventuale probabilità che ci sia una perdita di dati dovuta a fattori esogeni. Sotto questo punto di vista è logico intuire come la definizione di (ϵ, δ) -differential privacy, pur contemplando un panorama più ampio che di fatto trova una corrispondenza verosimile con la pluralità di scenari in cui un database può trovarsi, sia meno forte della definizione scritta nella versione che contempla solo il parametro ϵ . Al pari del parametro ϵ , neanche per il parametro δ esiste una legge scritta per l'assegnamento del suo valore. Tipicamente si sceglie δ in modo tale che sia inferiore rispetto alla quantità $\frac{1}{n}$, dove n rappresenta il numero di record del database. In questo modo, ogni persona rappresentata da un record ha una probabilità minima pari a δ che i suoi dati trapelino.

2.5 PROPRIETÀ DELLA DIFFERENTIAL PRIVACY

2.5.1 PRIVACY BUDGET

Per illustrare il concetto di privacy budget si può supporre di disporre di un database D supervisionato da un algoritmo K . Supponendo, inoltre, che le query che giungono al vaglio di K non siano così tanto specifiche da non riuscire a superare il suo controllo preliminare atto alla valutazione della loro specificità, K svolge il suo lavoro aggiungendo una certa quantità di rumore alle risposte fornite da D . Verosimilmente, è possibile anche ammettere che K non abbia ispezionato alcuna query per il momento. Stando così le cose, la privacy dei dati memorizzati in D è momentaneamente garantita al 100%.

Ogni volta che a D giunge una nuova query q_i da una sorgente S , esso risponde e ogni risposta compromette la privacy dei dati di un valore χ_i : dopo la prima query (q_1) la privacy dei dati è garantita al $100\% - \chi_1$, dopo la seconda query (q_2) la privacy dei dati è garantita al $100\% - \chi_1 - \chi_2$, ecc...

Se non esistesse alcuno stratagemma aggiuntivo oltre a quello implementato da K per la valutazione della specificità delle query, si arriverebbe più o meno facilmente alla condizione per cui la privacy dei dati sarebbe garantita allo 0%. In questa situazione i dati sarebbero privi di qualunque forma di protezione che ne regoli la visibilità.

Per garantire una significativa quantità di privacy dei dati, i proprietari di database possono imporre un tetto massimo per quanto riguarda il numero di query che possono essere sottoposte ai database stessi. Questo è ciò che sta dietro al concetto di *budget*.

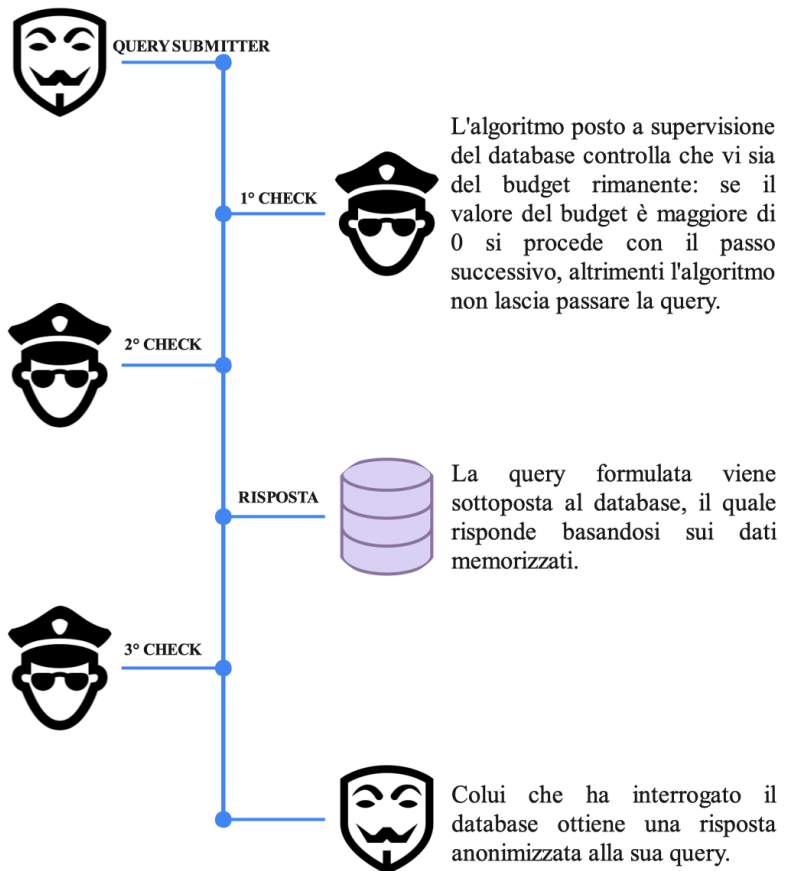
Formalmente, il budget è un numero che valorizza la massima quantità di privacy che può essere persa da un database. Se S non ha mai interrogato D , K assegna al parametro del budget associato a S il valore 1. Il budget è correlato mediante sottrazione alla funzione di valutazione della specificità delle query realizzata da K : più le query sono specifiche per l'insieme di dati immagazzinati in D e più la funzione di valutazione restituisce un valore elevato che, successivamente, deve essere sottratto al valore residuo del budget. Il procedimento è da iterare per ogni query formulata e il risultato derivante dalla sottrazione rappresenta il discriminante: fino a quando il valore aggiornato del budget è maggiore di 0, K permette a D di rispondere. Non appena il valore del budget assume valore 0, o un valore prossimo a 0 che lo renderebbe nullo o addirittura negativo qualunque sia la successiva query, K nega qualunque possibilità di interazione tra S e D .

2.5.1.1 SCHEMA RIASSUNTIVO

Qualcuno vuole interrogare il database per poterne estrarre delle informazioni. Per poter perseguire il suo scopo formula una query.

L'algoritmo posto a supervisione del database valuta la specificità della query. Calcolando la differenza tra il valore del budget rimanente e il valore calcolato dall'algoritmo riguardo la specificità della query, se si ottiene 0 o un valore negativo l'algoritmo non lascia passare la query, altrimenti si procede con il passo successivo.

L'algoritmo posto a supervisione del database anonimizza la risposta fornita dal database.



2.5.2 SENSITIVITÀ

Fino ad ora si è detto che l'algoritmo posto a supervisione del database aggiunge una certa quantità di rumore alle risposte fornite dal database stesso alle query che gli giungono. Giunti a questo punto, tuttavia, la domanda sorge spontanea: “quanto rumore deve essere aggiunto?”. A questa domanda si può rispondere facendo ricorso al concetto di sensitività di una query.

Considero due database D_1 e D_2 che differiscono per il solo record della persona di cui si vuole salvaguardare la privacy. Considero, inoltre, una interrogazione q ai due database che fornisca un risultato numerico che riguarda un certo numero n di attributi di una persona.

Indicando con Δq la cosiddetta sensitività della risposta, ovvero la differenza tra le risposte ottenute interrogando i due database, si ha che:

$$\Delta q = |q(D_1) - q(D_2)|$$

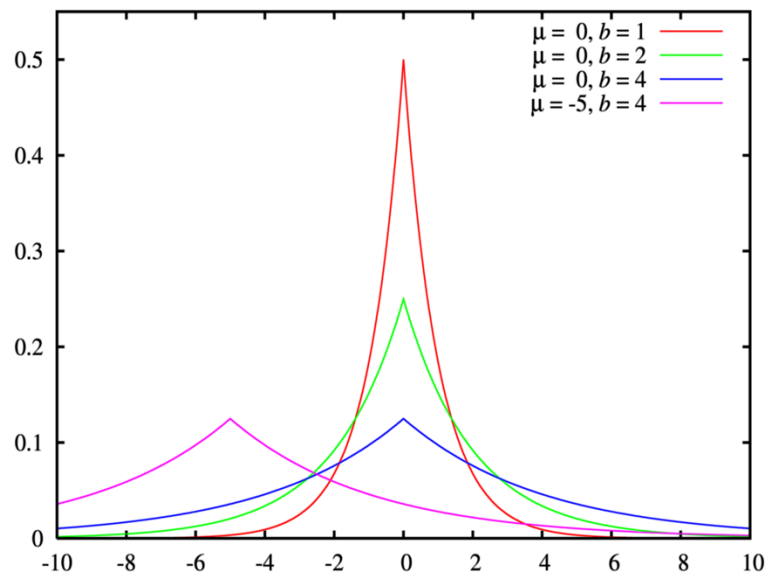
2.5.3 MECCANISMO DI LAPLACE

Oltre alla domanda “quanto rumore deve essere aggiunto?”, per poter implementare correttamente un meccanismo che garantisca la privacy dei dati immagazzinati in un database tramite la tecnica della privacy differenziale occorre chiedersi anche: “con quale criterio l’algoritmo posto a supervisione del database estrae il rumore casuale che deve essere aggiunto alle risposte fornite dal database?”.

Il meccanismo di Laplace comporta l’aggiunta ad ogni risposta di ogni query di un rumore casuale estratto dalla distribuzione di Laplace. La distribuzione di Laplace è una distribuzione simmetrica che presenta un valore medio μ che può essere sia positivo sia negativo e un parametro di scala $b = \frac{\Delta q}{\epsilon}$. In particolare, una variabile aleatoria Y ha distribuzione di Laplace $Y \sim \text{Lap}(\mu, b)$ se ha come funzione di densità di probabilità:

$$f(x|\mu, b) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}$$

Alcuni esempi di funzioni di densità di probabilità di Laplace sono:



Dall’immagine si può notare come il valore medio sia il valore più probabile (cioè la densità di probabilità raggiunge il suo massimo in questo punto) e man mano che ci si allontana da questo punto la probabilità decresce. In particolare:

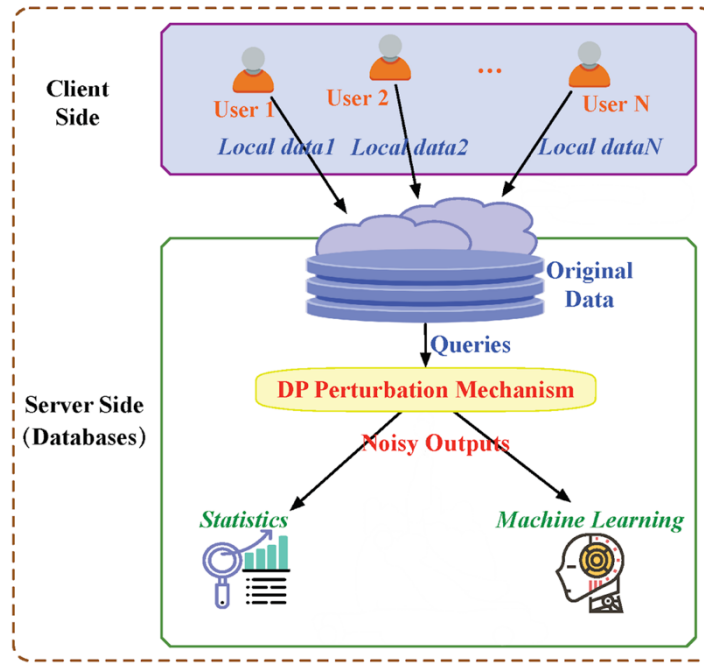
$$\mathbb{E}[Y] = \mu$$

$$\mathbb{V}[Y] = 2b^2$$

2.6 GDP, LDP, ESA

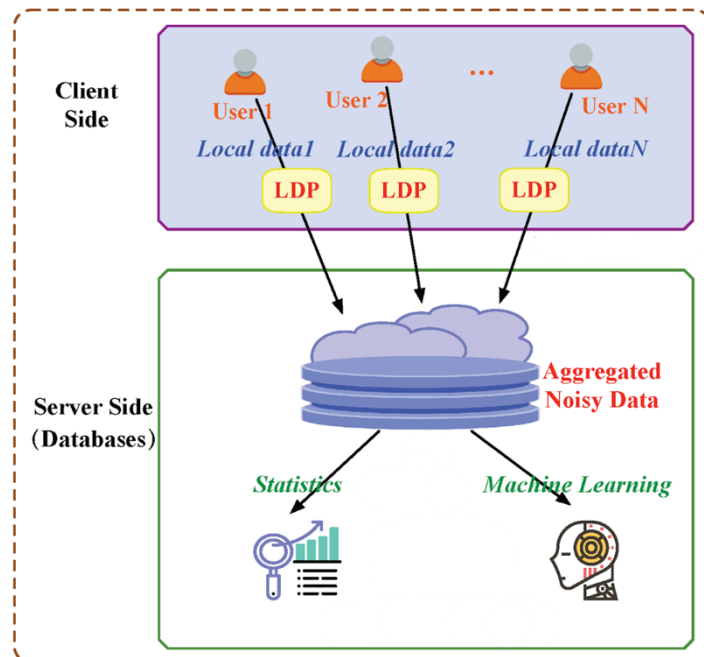
Quando si parla della tecnica della privacy differenziale si può fare riferimento a due diversi pattern implementativi:

- GDP, Global Differential Privacy \Rightarrow modello globale:



A General Processing Framework of Differential Privacy

- LDP, Local Differential Privacy \Rightarrow modello locale:

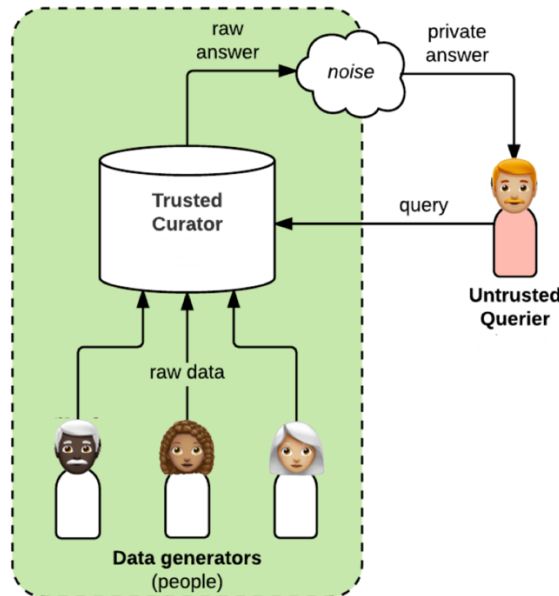


A General Processing Framework of Local Differential Privacy

Per capire la differenza tra GDP e LDP bisogna capire chi ha accesso ai dati reali.

2.6.1 GDP (GLOBAL DIFFERENTIAL PRIVACY)

Nel modello GDP è il trusted curator (aggregatore fidato), ovvero il proprietario di un database sul quale viene applicato il meccanismo GDP, ad avere l'accesso ai dati reali:



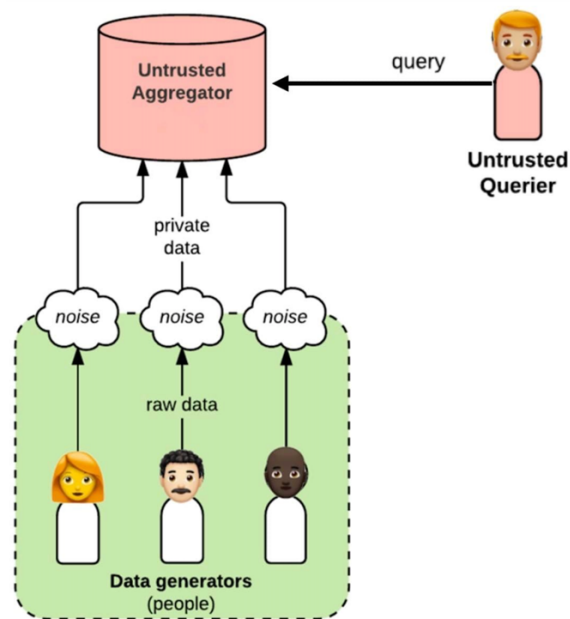
Come si può vedere, ogni utente manda i propri dati personali reali al trusted curator. È compito del trusted curator garantire un trattamento degli stessi che preveda una loro trasformazione mediante GDP. Il curator è trusted nel senso che è richiesto agli utenti che forniscono i dati reali una fiducia nei suoi confronti perché applichi correttamente il modello. In questo contesto, gli utenti che forniscono i dati reali sono i data generators.

In GDP il meccanismo di anonimizzazione viene applicato una volta sola nel momento in cui al database giungono interrogazioni da parte di utenti esterni, gli untrusted querier. A processo ultimato, l'aggregatore può condividere i dati anonimizzati con terzi.

Questo modello, per quanto sicuro contro gli untrusted querier, può non apparire sufficientemente sicuro agli occhi data generators. Il problema sta nel fatto che il curator ha necessità di conoscere i dati reali e non è sempre detto che il curator sia effettivamente trusted. Possono impersonificare questa figura compagnie non attendibili che puntano ad arricchirsi vendendo i dati reali oppure i governi di regimi autoritari per monitorare le attività dei singoli cittadini. Può anche essere vero che i data generators decidano autonomamente e consapevolmente di non fidarsi di terze parti per il trattamento dei loro dati. Inoltre, con il modello GDP i dati reali vengono immagazzinati in un solo luogo e questo aumenta le probabilità che il sistema possa essere, ad esempio, preso come bersaglio da parte di attacchi hacker.

2.6.2 LDP (LOCAL DIFFERENTIAL PRIVACY)

Nel modello LDP è contemplata la figura del untrusted aggregator (aggregatore non fidato), il quale, in quanto untrusted, non ha accesso ai dati reali. Il modello in questione prende piede dall'osservazione fatta nel paragrafo precedente riguardo il fatto che i data generators possono non fidarsi del curator per un qualsiasi lecito motivo. Secondo l'approccio LDP, quindi, i data generators, decidendo a priori di non fidarsi di nessuno per il trattamento dei propri dati privati, li devono anonimizzare preventivamente:



Dopo aver collezionato i dati rumorosi, l'aggregatore può elaborare varie statistiche ed eventualmente pubblicare i propri risultati. In linea di principio, sarebbe possibile per lui pubblicare l'intero database in virtù del fatto che i dati che gli giungono sono già anonimi. Il grande vantaggio del modello LDP è che i data generators non si devono fidare di una terza figura della quale sono ignote le intenzioni. Sull'altro fronte, tuttavia, il modello ha come svantaggio una pesante politica di anonimizzazione. Come già detto, nel modello GDP l'anonymizzazione viene effettuata sull'intero database nel momento in cui questo viene interrogato. Nel modello LDP, invece, si ha una anonimizzazione per ogni singolo dato che viene aggiunto alla base di dati. Questo fa sì che le statistiche estraibili da un database sul quale è implementato un meccanismo LDP siano meno attendibili di quelle che verrebbero estratte dal medesimo database se su questo venisse implementato un meccanismo GDP.

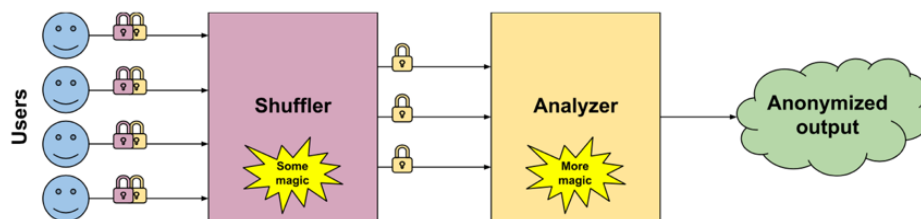
LDP è un meccanismo implementato da diverse società, tra le quali vale la pena citare Google con RAPPOR e Apple con Private Count Mean Sketch.

2.6.3 ESA (ENCODE, SHUFFLE, ANALYZE)

Fino a poco tempo fa non c'era una via di mezzo tra le due opzioni GDP e LDP. La scelta implementativa, in un certo senso, era binaria: GDP o LDP.

L'infrastruttura ESA si pone come obiettivo quello di trovare un giusto compromesso tra la pesante politica di anonimizzazione dei dati tipica del modello LDP e la totale fiducia richiesta agli utenti verso la figura del trusted curator nel modello GDP.

Il modello ESA, a differenza dei suoi due predecessori che contemplano le due figure dell'utente e dell'aggregatore, ne prevede tre: encoder, shuffler e analyzer.



Encoder è un termine tecnico proprio di ESA che sta ad indicare l'utente, ovvero quell'entità da cui provengono i dati reali da anonimizzare. L'encoder ha tre compiti:

1. Anonimizzare i dati (in tale fase si può notare l'impronta data dal modello LDP).
2. Codificare i dati anonimizzati su due differenti livelli di cifratura.
3. Inoltrare i dati anonimizzati e cifrati allo shuffler.

Lo shuffler è un processo intermedio tra l'encoder e l'analyzer. Anch'esso ha tre compiti:

1. Decriptare il primo dei due livelli di cifratura imposti dell'encoder. A seguito di ciò, lo shuffler trova due parametri: user ID e group ID. User ID è un parametro proprio di ogni dato, mentre group ID è un parametro che descrive il tipo del dato. Una precisazione è doverosa: né user ID né group ID contengono il valore del dato anonimizzato dato che questo è ancora protetto da un livello di cifratura.
2. Dati differenti che presentano lo stesso valore per group ID sono da indentificare come appartenenti alla stessa categoria, dunque lo shuffler, sfruttando il group ID, raggruppa tutti i dati che presentano lo stesso valore di questo parametro e li conta.
3. Se i gruppi sono abbastanza numerosi da poterci elaborare delle statistiche attendibili, li inoltra al analyzer, altrimenti li conserva in attesa che aumentino.

L'analyzer per ESA è l'analogo del trusted curator per il modello GDP. Ha due compiti:

1. Decriptare il secondo livello di cifratura imposto dall'encoder, giungendo così ai dati rumorosi privi di ogni forma di criptazione.
2. Analisi statistiche, machine learning, ...

CAPITOLO 3: SETUP ENVIRONMENT

3.1 PRESENTAZIONE DEI TOOLS

Nella fase operativa è stato necessario sfruttare i seguenti tools:

- Git, un software che ha permesso di scaricare in locale da GitHub i files della libreria fornita da Google.
- Bazel, un software che ha permesso la compilazione di tutto il codice presente nella libreria fornita da Google.

3.2 DESCRIZIONE DEL PROBLEMA

Nella fattoria di Fred ci sono 182 animali. Ogni giorno Fred li nutre con delle carote e ognuno degli animali tiene il conto di quante carote ha mangiato nell'arco della giornata. Giunta sera, Fred interroga tutti gli animali della fattoria chiedendogli, ad esempio, quante carote hanno mangiato in totale oppure chi è l'animale che ne ha mangiate di più. Gli animali, timorosi che Fred possa usare le informazioni che gli fornirebbero contro di loro, decidono di mentirgli camuffando i dati reali.

3.3 DALLA TEORIA ALLA PRATICA

La trattazione successiva prevede l'applicazione della teoria esposta nel capitolo 2 al problema della fattoria di Fred. Nella fattispecie, i macro-passi che verranno seguiti sono:

- Fissare un valore di ϵ sfruttando tre valori rappresentativi:
 - $\epsilon = 0,01$;
 - $\epsilon = 1$;
 - $\epsilon = 100$.
- Per ciascun valore di ϵ verrà mostrato uno snapshot rappresentante una simulazione che ritrae una delle infinite possibili situazioni in cui si può venire a trovare Fred nel momento in cui interroga gli animali della fattoria.
- Per ciascuna query presente nella simulazione verrà presentata una breve discussione sui risultati ottenuti e, contestualmente, verrà mostrato un grafico che mostra l'andamento dei valori delle risposte sulla base di altre 20 simulazioni.

3.3.1 $\epsilon = 0,01$

Simulazione:

```
It is a new day. Farmer Fred is ready to ask the animals about their carrot consumption.
-----
Farmer Fred asks the animals how many total carrots they have eaten. The animals know the true sum but report the differentially private sum to Farmer Fred.
But first, they ensure that Farmer Fred still has privacy budget left.

Privacy budget remaining: 1.00
True sum: 9649
DP sum: 314726
-----
Farmer Fred catches on that the animals are giving him DP results. He asks for the mean number of carrots eaten, but this time, he wants some additional
accuracy information to build his intuition.

Privacy budget remaining: 0.75
True mean: 53.02
The animals were not able to get the private mean with the current privacy parameters. This is due to the small size of the dataset and random chance.
-----
Fred wonders how many gluttons are in his zoo. How many animals ate over 90 carrots?

Privacy budget remaining: 0.50
True count: 21
DP count: -192
-----
'And how gluttonous is the biggest glutton of them all?' Fred exclaims. He asks how many carrots the animal that ate the most has eaten.

Privacy budget remaining: 0.25
True max: 100
DP max: 138
-----
Fred also wonders how many animals are not eating any carrots at all.

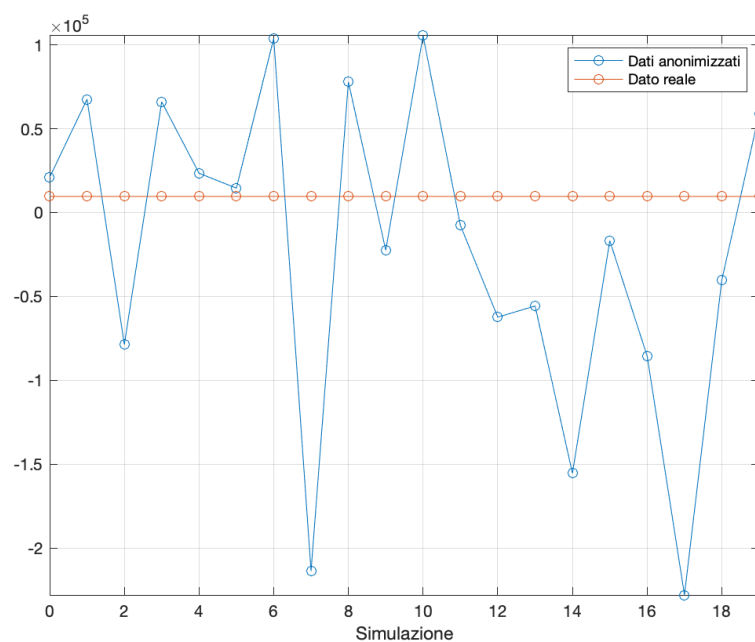
Privacy budget remaining: 0.00
Error querying for count: Not enough privacy budget.
The animals notice that the privacy budget is depleted. They refuse to answer any more of Fred's questions for risk of violating privacy.
```

Analisi della prima query.

Il risultato anonimizzato è fuori da ogni logica, ma ciò non deve stupire: per valori di $\epsilon \rightarrow 0$ si ha una completa garanzia di riservatezza dei dati a discapito di una informazione estraibile a fini statistici pressoché nulla. Contestualizzando il dato fornito dalla specifica simulazione, durante l'arco della giornata ogni animale dovrebbe aver mangiato

$$\frac{314726}{182} \approx 1729 \text{ carote.}$$

Grafico:



Analisi della seconda query.

Gli animali si rifiutano rispondere alla query perché sennò farebbero capire a Fred che hanno mentito alla sua precedente interrogazione. In tal caso, potrebbero venire compromesse le future porzioni di carote.

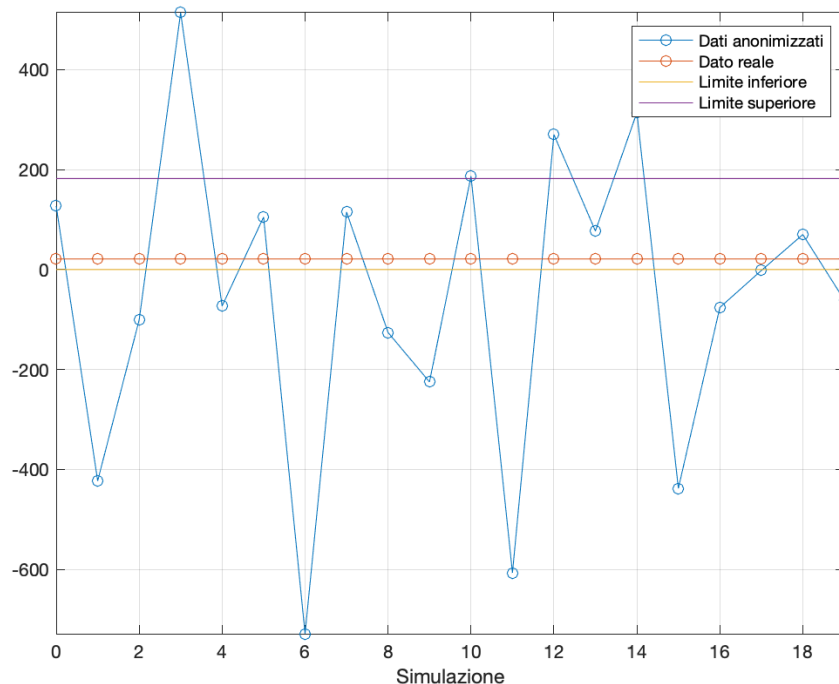
Il motivo di questo comportamento è da ricercare nel fatto che la query sottoposta da Fred agli animali è troppo specifica sia rispetto alla quantità limitata di animali presenti nella fattoria sia rispetto al valore di ε .

In questo caso il grafico è inutile perché sia le dimensioni del database sia il valore di ε non variano.

Analisi della terza query.

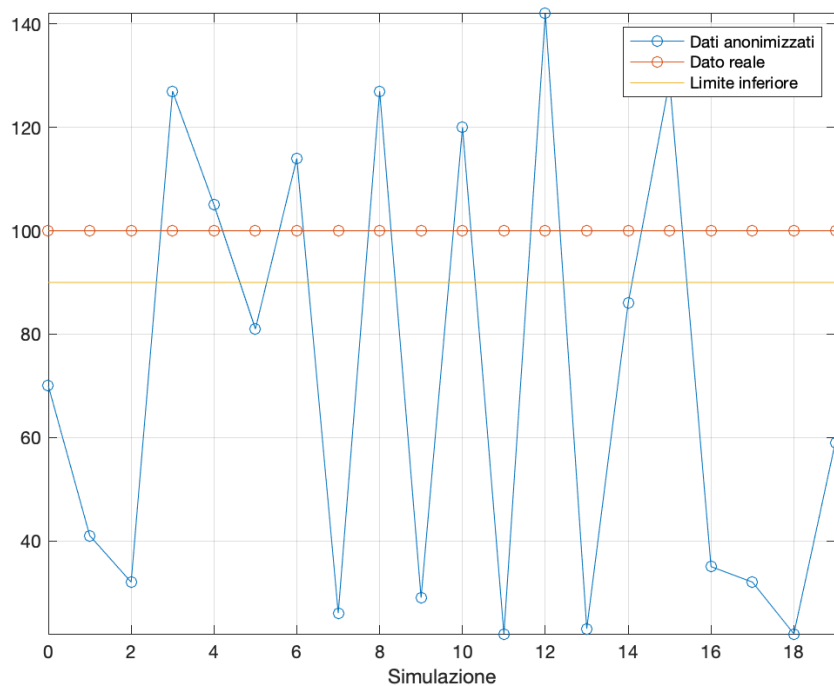
Il commento è perfettamente in linea con quanto riportato nei confronti della prima query. Peraltro, risulta essere negativo un valore che deve essere necessariamente compreso nell'intervallo $0 \leq x \leq 182$.

Grafico:



Analisi della quarta query.

In generale, il risultato anonimizzato ottenuto in risposta alla quarta query deve avere senso se paragonato con il rispettivo valore reale e deve rispecchiare quanto messo in evidenza con la risposta alla terza query. Stando alla terza query, infatti, dato che la risposta reale è un numero maggiore di 0 (il valore anonimizzato della terza query, come già detto, è privo di ogni fondamento matematico e utilità a fini statistici), ovvero c'è almeno un animale che ha mangiato almeno 90 carote, è logico aspettarsi che entrambe le risposte alla quarta query siano dei numeri strettamente maggiori di 90. La risposta reale alla quarta query, ovvero il numero 100, rispetta questa condizione, tuttavia, sebbene in questa specifica simulazione si sia ottenuto un valore pari a 138 per la risposta anonimizzata, e dunque coerente con quanto appena scritto, il valore di ϵ fa sì che nelle successive simulazioni possano essere ottenuti anche valori minori di 90, come si evince dal grafico sottostante:



3.3.2 $\epsilon = 1$

Simulazione:

```
It is a new day. Farmer Fred is ready to ask the animals about their carrot consumption.
-----
Farmer Fred asks the animals how many total carrots they have eaten. The animals know the true sum but report the differentially private sum to Farmer Fred.
But first, they ensure that Farmer Fred still has privacy budget left.

Privacy budget remaining: 1.00
True sum: 9649
DP sum: 9397
-----
Farmer Fred catches on that the animals are giving him DP results. He asks for the mean number of carrots eaten, but this time, he wants some additional
accuracy information to build his intuition.

Privacy budget remaining: 0.75
True mean: 53.02
The animals were not able to get the private mean with the current privacy parameters. This is due to the small size of the dataset and random chance.
-----
Fred wonders how many gluttons are in his zoo. How many animals ate over 90 carrots?

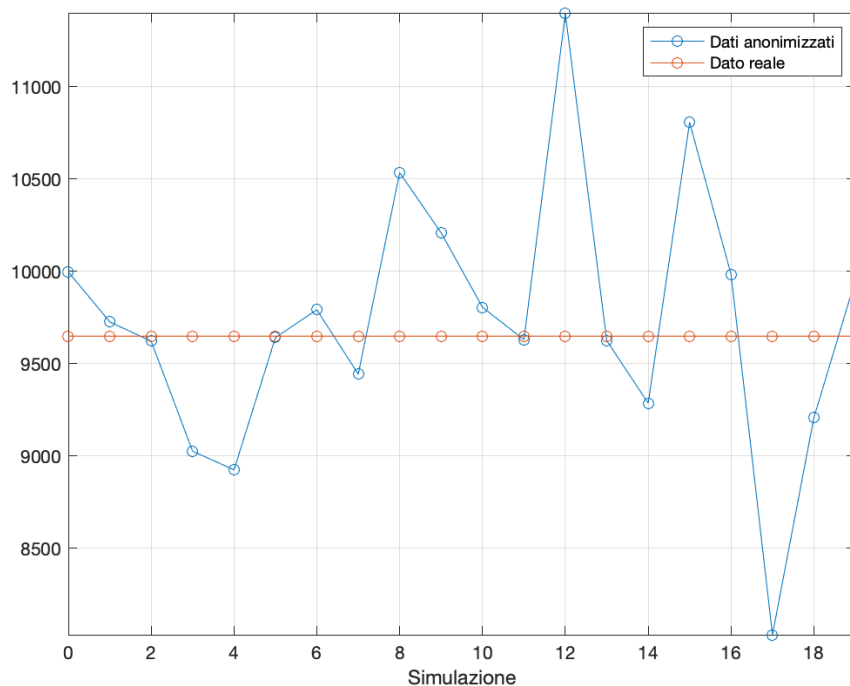
Privacy budget remaining: 0.50
True count: 21
DP count: 26
-----
'And how gluttonous is the biggest glutton of them all?' Fred exclaims. He asks how many carrots the animal that ate the most has eaten.

Privacy budget remaining: 0.25
True max: 100
DP max: 74
-----
Fred also wonders how many animals are not eating any carrots at all.

Privacy budget remaining: 0.00
Error querying for count: Not enough privacy budget.
The animals notice that the privacy budget is depleted. They refuse to answer any more of Fred's questions for risk of violating privacy.
```

Analisi della prima query.

Il risultato ha senso e il grafico sottostante lo conferma:

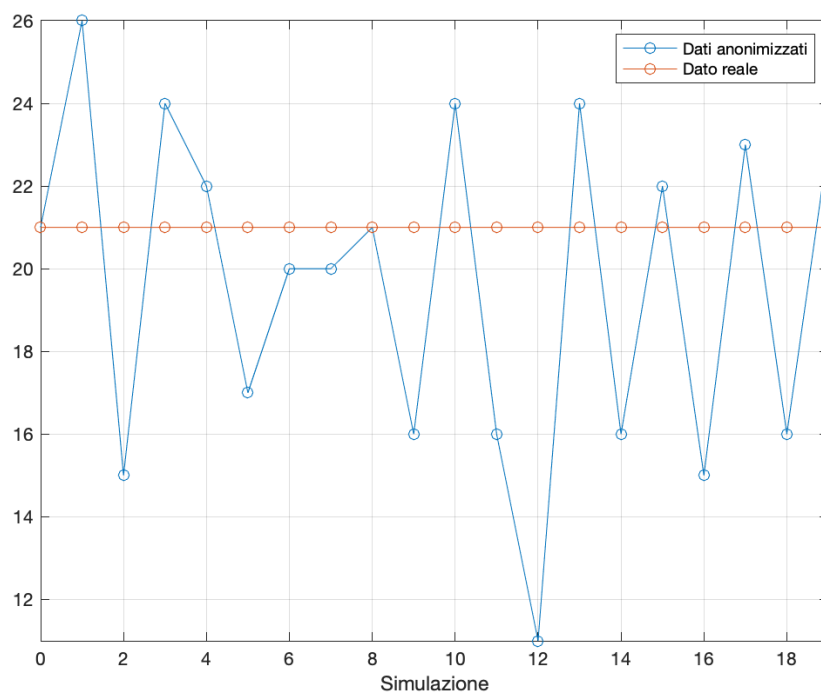


Analisi della seconda query.

Il commento è del tutto analogo a quello riportato nel caso per cui si è fissato $\epsilon = 0,01$.

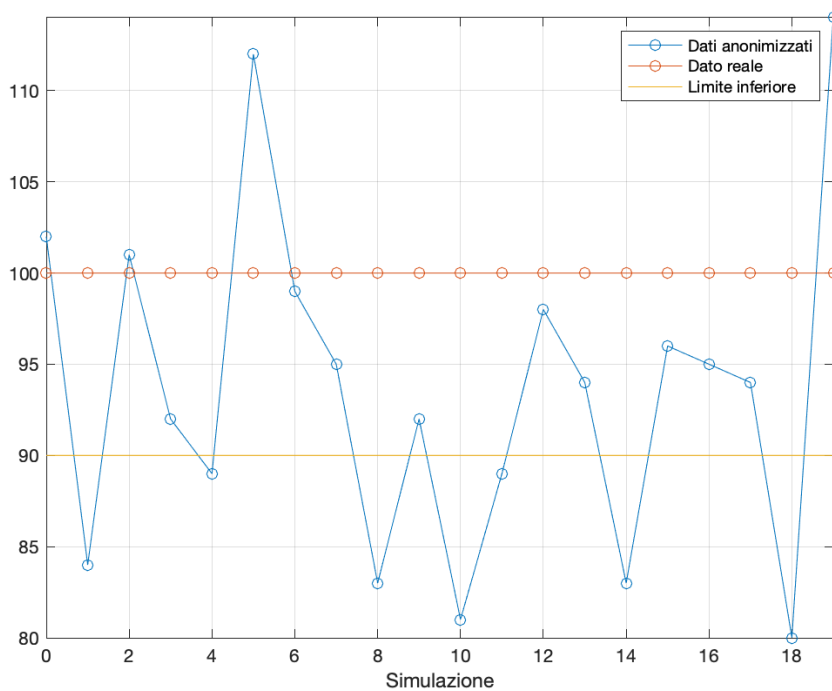
Analisi della terza query.

Il risultato ha senso e il grafico sottostante lo conferma:



Analisi della quarta query.

Il commento è perfettamente in linea con quanto riportato nei confronti della quarta query nel caso per cui $\epsilon = 0,01$. Il valore 74, essendo incoerente con la risposta fornita dagli animali alla terza query, darebbe a Fred la conferma che gli animali gli hanno mentito.



3.3.3 $\epsilon = 100$

Simulazione:

```
It is a new day. Farmer Fred is ready to ask the animals about their carrot consumption.
-----
Farmer Fred asks the animals how many total carrots they have eaten. The animals know the true sum but report the differentially private sum to Farmer Fred.
But first, they ensure that Farmer Fred still has privacy budget left.

Privacy budget remaining: 1.00
True sum: 9649
DP sum: 9650
-----
Farmer Fred catches on that the animals are giving him DP results. He asks for the mean number of carrots eaten, but this time, he wants some additional
accuracy information to build his intuition.

Privacy budget remaining: 0.75
True mean: 53.02
DP mean: 53.04
-----
Fred wonders how many gluttons are in his zoo. How many animals ate over 90 carrots?

Privacy budget remaining: 0.50
True count: 21
DP count: 21
-----
'And how gluttonous is the biggest glutton of them all?' Fred exclaims. He asks how many carrots the animal that ate the most has eaten.

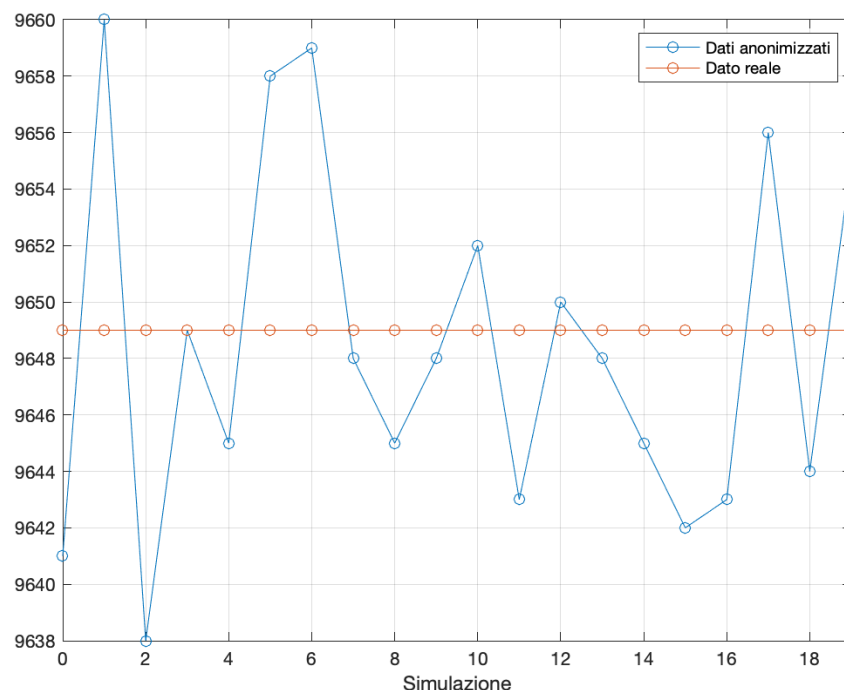
Privacy budget remaining: 0.25
True max: 100
DP max: 88
-----
Fred also wonders how many animals are not eating any carrots at all.

Privacy budget remaining: 0.00
Error querying for count: Not enough privacy budget.
The animals notice that the privacy budget is depleted. They refuse to answer any more of Fred's questions for risk of violating privacy.
```

Analisi della prima query.

Il valore reale e quello anonimizzato sono talmente simili da poter essere considerati indistinguibili, ma questo non deve stupire: per valori di $\epsilon \rightarrow \infty$ si ha una informazione derivante da un dato anonimizzato che è estremamente verosimile con quella estraibile dal dato reale. Tutto ciò avviene senza la benché minima garanzia di riservatezza dei dati.

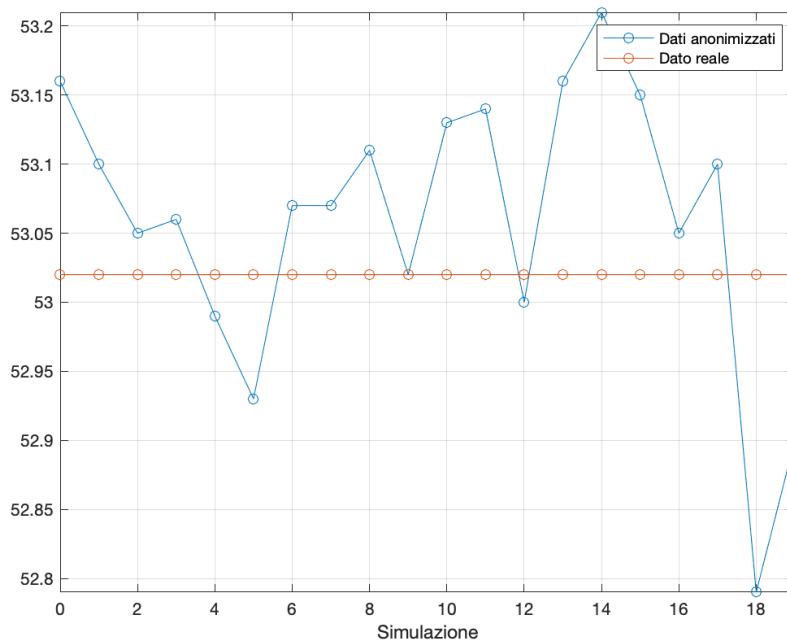
Grafico:



Analisi della seconda query.

In questo caso il valore di ε è tale da far sì che gli animali rispondano, seppur le dimensioni del database non siano cambiate. Il motivo è il medesimo di quello riportato per la prima query.

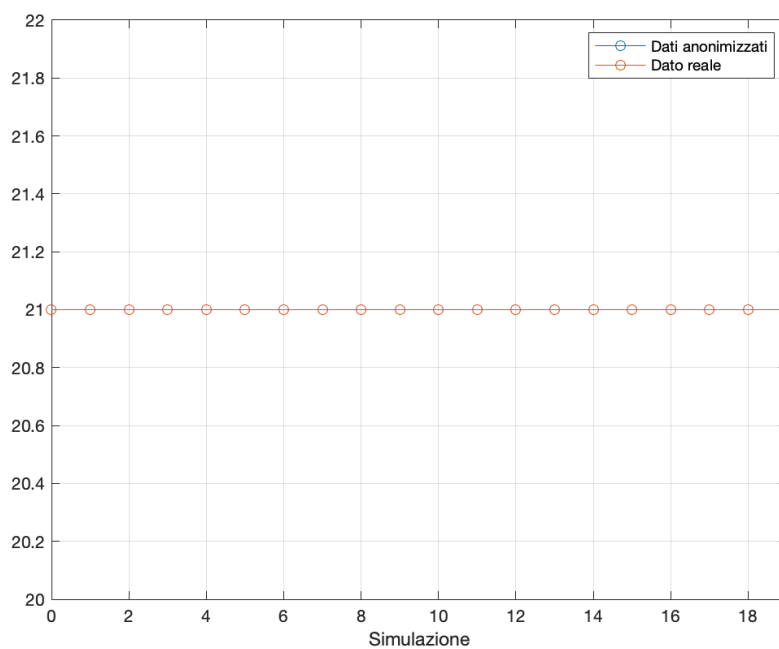
Grafico:



Analisi della terza query.

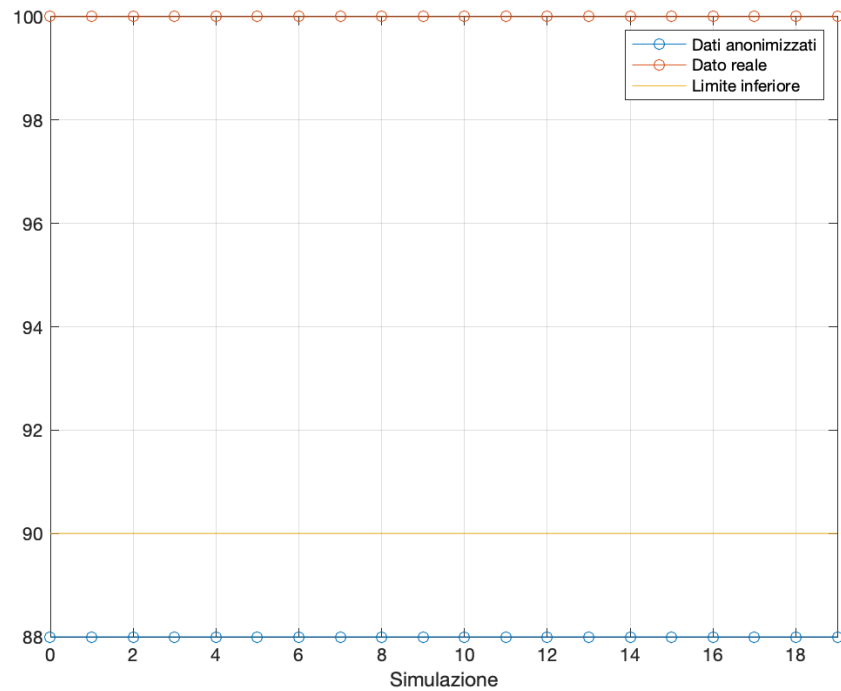
Il commento è perfettamente in linea con quello riportato nei confronti della prima query.

Grafico:



Analisi della quarta query.

Il commento è perfettamente in linea con quanto riportato nei confronti della quarta query per entrambi i casi $\epsilon = 0,01$ e $\epsilon = 1$. Il valore 88, essendo incoerente con la risposta fornita dagli animali alla terza query, darebbe a Fred la conferma che gli animali gli hanno mentito.



3.4 OSSERVAZIONI

Ciò che segue sono due osservazioni volte a rendere ancora più chiare le situazioni ritratte dalle simulazioni effettuate al variare del parametro ε :

- Il caso della libreria fornita da Google rappresenta una semplificazione in quanto viene meno la logica della funzione implementata dall'algoritmo posto a supervisione del database relativa alla valutazione della specificità di ogni query. Stando a quel ragionamento, la funzione appena citata avrebbe dovuto calcolare un valore da sottrarre al budget che fosse direttamente proporzionale alla specificità di ciascuna query formulata da Fred. Tutto ciò è stato sostituito da una sottrazione dal valore del budget rimanente per un valore deciso a priori e uguale per ogni query formulata, ovvero 25%. In altre parole, è stato deciso a priori che le query debbano avere un peso del 25% nei confronti del budget rimanente.
- In nessun caso è mai stata discussa la risposta degli animali della fattoria alla quinta query formulata da Fred. La motivazione è semplice: in accordo con quanto riportato sia nel paragrafo 2.5.1 sia al punto precedente, Fred ha a disposizione un budget viene decrementato del 25% per ogni query che decide di sottoporre agli animali. È logico, quindi, che dopo quattro query a cui gli animali rispondono, il budget sia totalmente consumato. Dato che Fred non ha più budget a disposizione, gli animali si rifiutano di rispondere.

BIBLIOGRAFIA

- [1]. Cynthia Dwork, Aaron Roth, *The Algorithmic Foundations of Differential Privacy*, Foundations and Trends in Theoretical Computer Science, vol. 9 (2014).
- [2]. Giuseppe D'Acquisto, Maurizio Naldi, *Big Data and Privacy by Design*, I diritti nella “rete” della rete, Anonimizzazione Pseudonimizzazione Sicurezza, Prefazione di Franco Pizzetti, G. Giappichelli Editore (2017).

SITOGRAFIA

- [1]. *Privacy differenziale - Differential privacy*,
https://it.abcdef.wiki/wiki/Differential_privacy
- [2]. Zanichelli editore, *La privacy #1*. Come e quando nasce il diritto alla privacy,
<https://www.youtube.com/watch?v=jDLkHurCpKQ>
- [3]. Zanichelli editore, *La privacy #2*. Il GDPR e il trattamento dei dati personali,
<https://www.youtube.com/watch?v=WUyBmkeLWxM>
- [4]. Zanichelli editore, *La privacy #3*. Le regole del GDPR,
<https://www.youtube.com/watch?v=LmCLJyEc4-g>
- [5]. *Guida completa al CCPA*, <https://www.iubenda.com/it/help/19153-guida-ccpa>
- [6]. *Che cos'è la LGPD*, <https://www.iubenda.com/it/help/26708-guida-lgpd>
- [7]. Provino A., *Privacy vs Utility Trade-off*, <https://andreaprovino.it/privacy-vs-utility-trade-off>
- [8]. *Netflix Prize*, <http://www.mathisintheair.org/wp/2017/03/come-netflix-capisce-il-film-che-volete-vedere-matematica-machine-learning-e-serie-tv/>
- [9]. *Differential Privacy Definition*, <https://medium.com/@shaistha24/differential-privacy-definition-bbd638106242>
- [10]. *Query “sensitivity” types and effects on differential privacy mechanism*,
<https://becominghuman.ai/query-sensitivity-types-and-effects-on-differential-privacy-mechanism-c94fd14b9837>
- [11]. Distribuzione di Laplace, https://it.wikipedia.org/wiki/Distribuzione_di_Laplace
- [12]. Desfontaines D., Local vs global differential privacy,
<https://desfontain.es/privacy/local-global-differential-privacy.html>
- [13]. Libreria di Google, <https://github.com/google/differential-privacy>