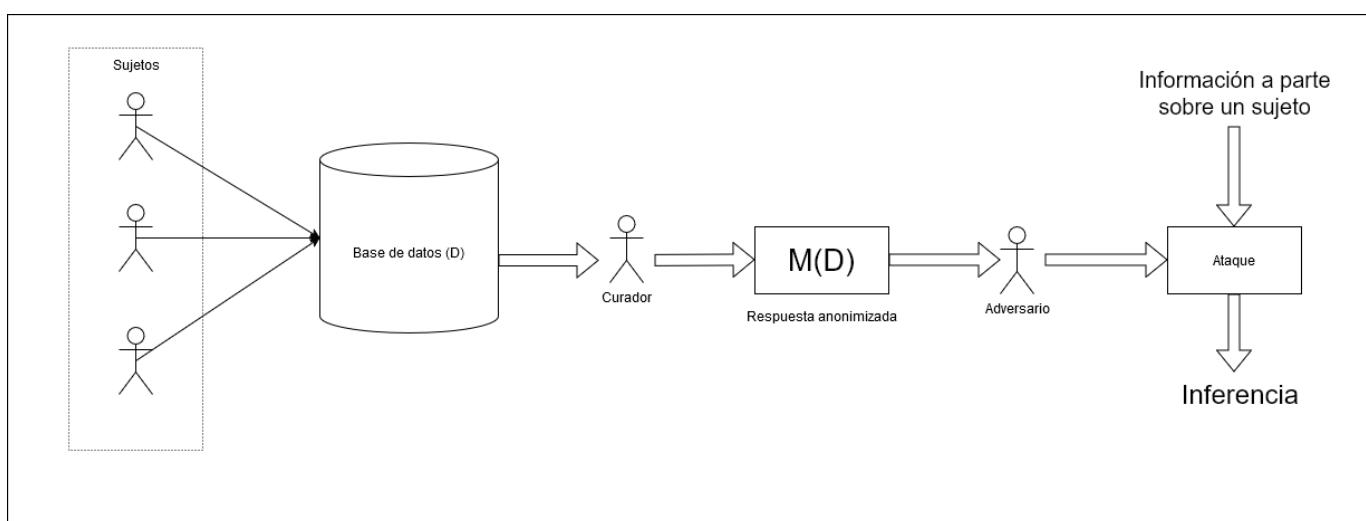


# [PAN] Privacidad Diferencial (Resumen)

## Caso base

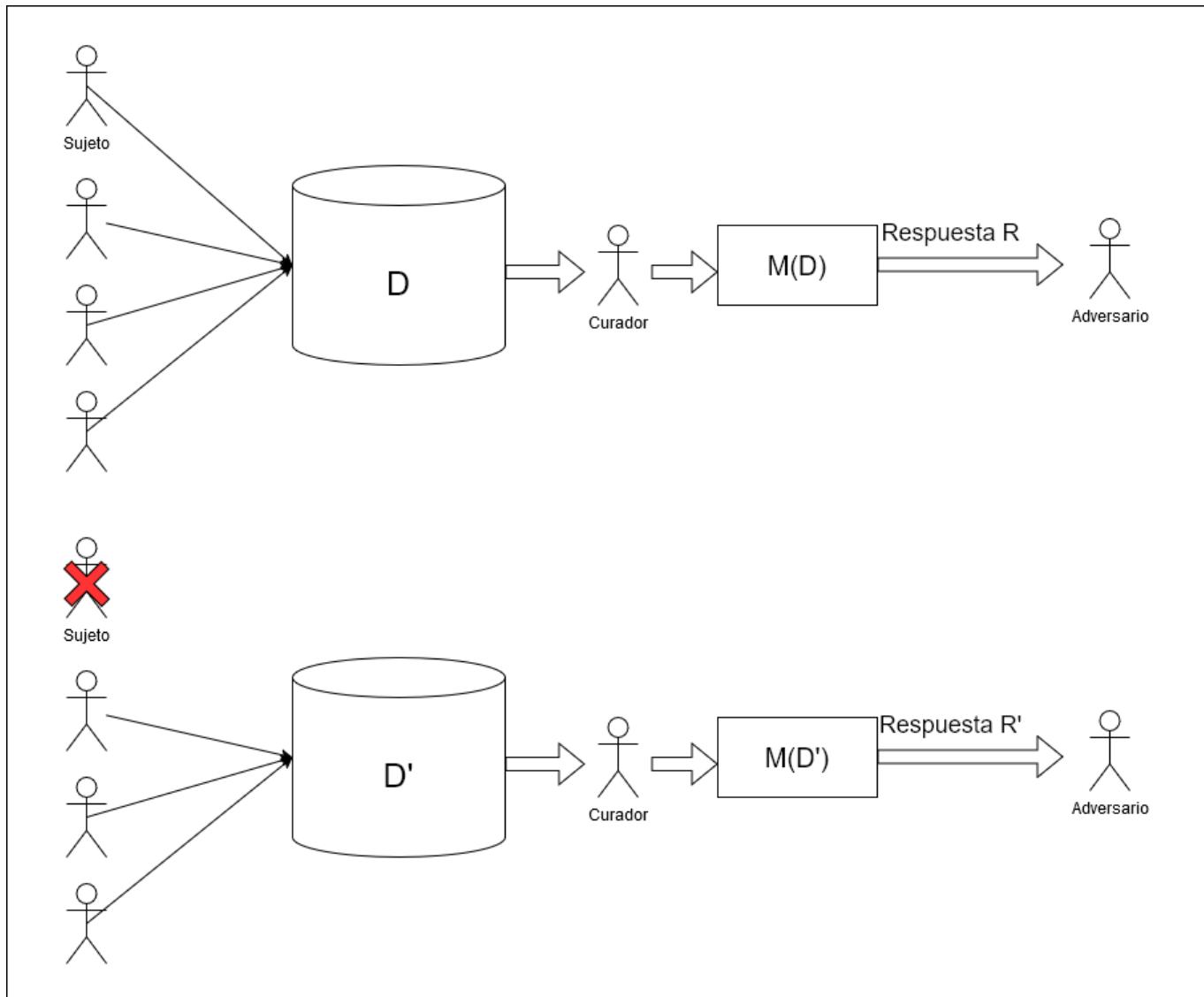
Tenemos un dataset  $D$  que contiene datos de usuarios, siendo cada fila los datos de un usuario. El Curador, que es una entidad de confianza para los usuarios, publica algunos datos usando un mecanismo  $M$  que da como resultado  $R=M(D)$ . El adversario trata de realizar inferencias sobre los datos  $D$  contenidos en  $R$ .

## De que protege la privacidad diferencial



La privacidad diferencial protege contra el riesgo del conocimiento de información sobre un sujeto usando inferencia con información obtenida a parte. De esta forma, observando la respuesta  $R$  no se puede cambiar lo que el adversario puede saber.

La clave para dificultar a un adversario identificar datos sobre un sujeto es poder crear dos salidas  $R=M(D)$  y  $R=M(D')$ , siendo  $D$  y  $D'$  dos datasets diferenciados por que el primero contiene al sujeto en cuestión y el segundo no, las cuales no puedan ser distinguibles la una de la otra. Para hacer esto se diseña el mecanismo  $M$ , el cual no puede ser determinístico, si no probabilístico.



La distribución de los datasets debe ser similar, es decir, dada una probabilidad  $R$  de que un dato viene del dataset  $D$ , esta tiene que ser similar a la probabilidad de que un dato venga del dataset  $D'$ . Los datasets que difieren en una fila son conocidos como vecinos. En resumidas cuentas, la probabilidad de que  $M(D)=R$  debe ser muy similar a la de que  $M(D')=R$ .

## Como definir distribuciones similares

**Definición tentativa de privacidad con parámetro P:** Un mecanismo  $M$  es privado si para todas las salidas posibles de  $R$  un todos los pares de datasets vecinos ( $D, D'$ ):

$$\Pr(M(D')=R) - P < \Pr(M(D)=R) < \Pr(M(D')=R) + P$$

From:  
<http://www.knoppia.net/> - Knoppia

Permanent link:  
[http://www.knoppia.net/doku.php?id=pan:res\\_privacidad\\_diferencial&rev=1736272346](http://www.knoppia.net/doku.php?id=pan:res_privacidad_diferencial&rev=1736272346)

Last update: 2025/01/07 17:52



