Data Anonymization

Need for privacy:
In the current scenario data is something very valuable similar to intellectual property.
Theft of data is also becoming more prevalent, In the present developing trends data is more valuable than money to a lot of organisations.
To prevent the misuse of data, theft of any valuable knowledge,to make the identities of the clients remain safe,event the transactions made , etc
So to avoid any illegal or mis use of the data that is obtained through any source permitted or non permitted to remain a liability and cause any heavy damage or loss we need privacy measures,

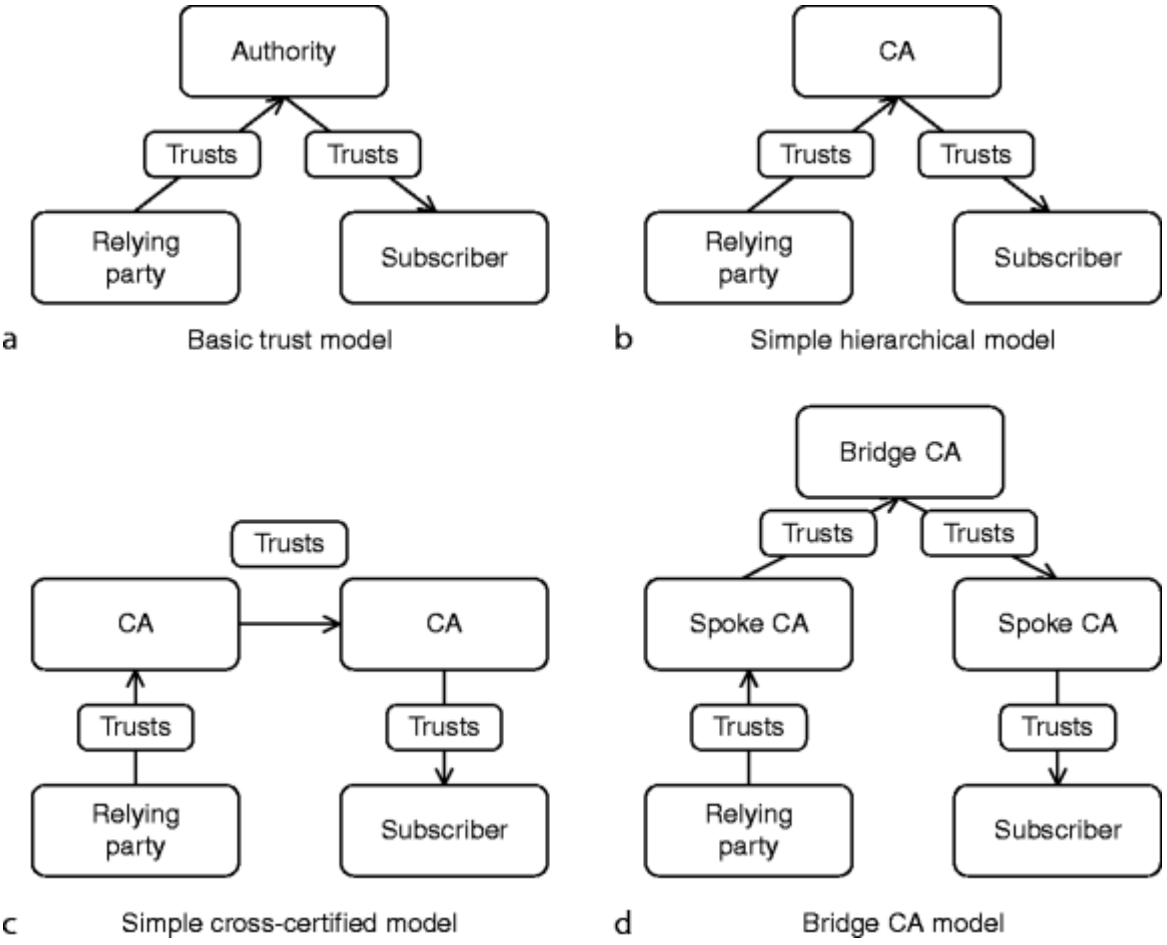
Challenges in applying privacy preservation:
Individualization of the data
Correlation between the variables/ attributes
Inferences delivered by the data with or without ML implementation\

K-Anonymity: attributes are suppressed or generalized until each row is identical with at least k-1 other rows. At this point the database is said to be k-anonymous. K-Anonymity thus prevents definite database linkages. At worst, the data released narrows down an individual entry to a group of k individuals

Tool used is Amnesia (online platform)
Amnesia is one of the many tools used for the anonymization process. Amnesia is also a flexible data anonymization tool that allows the removal of identifying information from data. Amnesia does not only remove direct identifiers like names, SSNs, etc., but also transforms secondary identifiers like birth date and zip code so that individuals cannot be identified in the data.

User Behaviour model

Trust models:
A trust Model is a collection of rules that informs applications on how to decide legitimacy of a Digital Certificate.



Privacy Preservation

IPFS:

The InterPlanetary File System (IPFS) is a protocol and peer-to-peer network for storing and sharing data in a distributed file system. IPFS uses content-addressing to uniquely identify each file in a global namespace connecting all computing devices.

IPFS allows users to not only receive but host content, in a similar manner to BitTorrent. As opposed to a centrally located server, IPFS is built around a decentralized system of user-operators who hold a portion of the overall data, creating a resilient system of file storage and sharing. Any user in the network can serve a file by its content address, and other peers in the network can find and request that content from any node who has it using a distributed hash table (DHT).