Shared Data Store

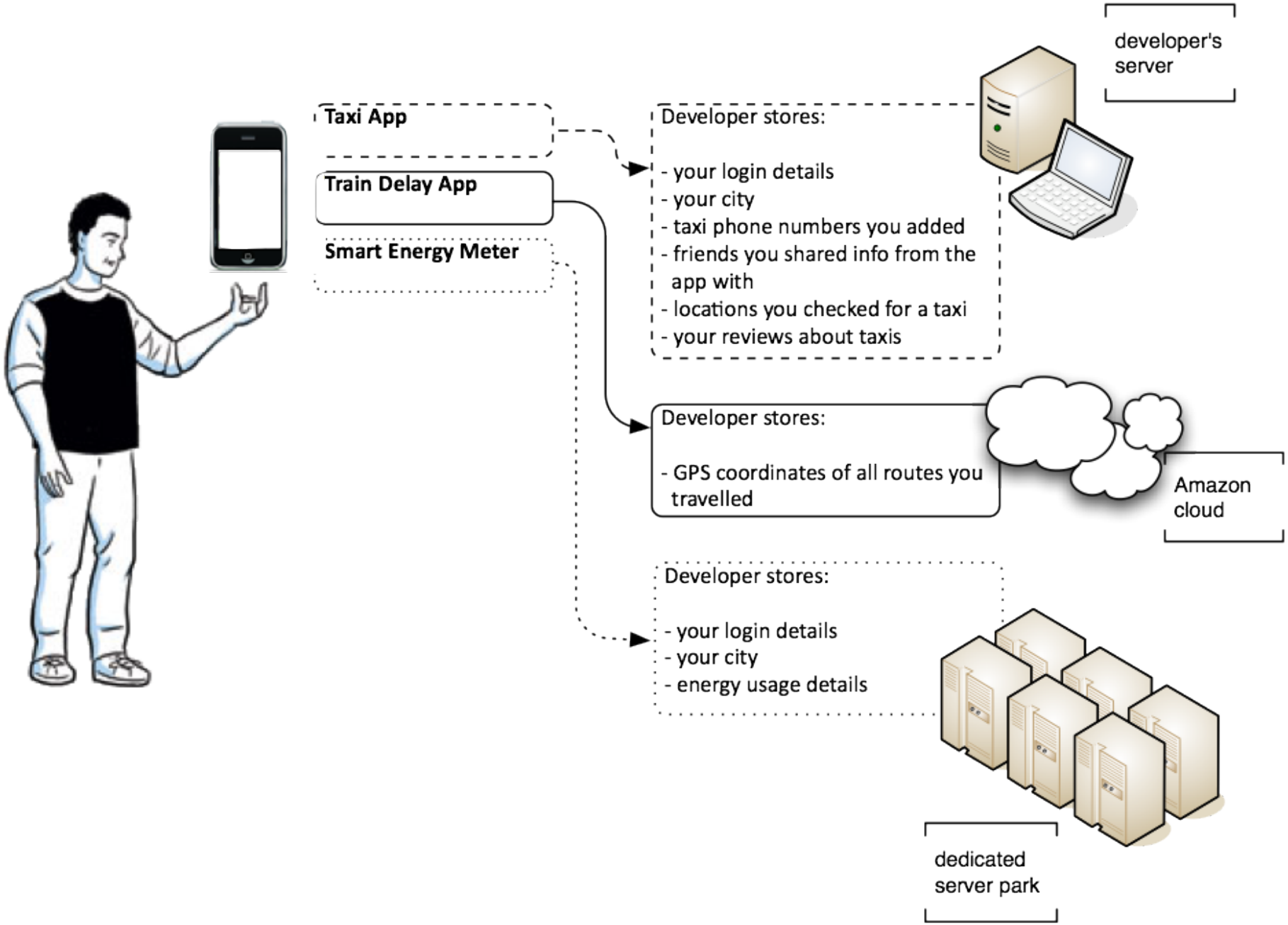
Tom Demeyer, tom@waag.org
Taco van Dijk, taco@waag.org

Shared Data Store (SDS)

- De afgelopen jaren is de hoeveelheid slimme applicaties en de gebruikers die er toegang toe hebben enorm toegenomen. Niet alleen zijn we langer en vaker online, we gebruiken ook meer verschillende tools die ons helpen om sneller en efficiënter met elkaar te communiceren, te werken en bijvoorbeeld te reizen.
- Met de toename van tools en applicaties groeit ook de hoeveelheid persoonlijke gegevens die we middels deze applicaties opslaan. De gemiddelde smartphone heeft meer dan 20 ingebouwde sensoren die informatie bijhouden over je bewegingsritme, de afstanden die je reist, de snelheid waarmee je beweegt etc. En wanneer je als gebruiker apps toegang tot deze informatie geeft, worden al die gegevens door externe partijen bijgehouden en opgeslagen.

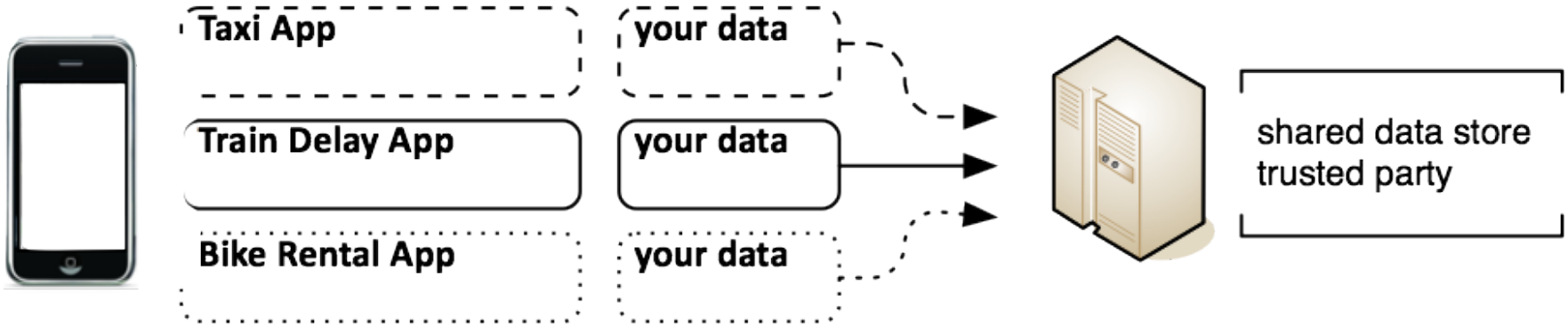
Shared Data Store (SDS)

- Vertrouwen wij als burgers al die duizenden bedrijven dat ze veilig met onze data omgaan? Dat niet ineens mijn logingegevens en creditcard informatie op straat liggen? Sommige partijen zijn betrouwbaarder dan anderen - hierbij speelt persoonlijke voorkeur een belangrijke rol. Alleen als gebruiker heb je momenteel in veel gevallen geen keuze; geen keuze in welke data je wilt delen en geen keuze waar je die data wilt opslaan.
- In het kader van het project Nederland Opent Data heeft Waag Society de 'shared data store' ontwikkeld. Het project draagt bij aan de bestaande vendor relationship management discussie en biedt een oplossing voor ontwikkelaars en burgers.


Taxi App
Train Delay App
Smart Energy Meter
developer's server
Developer stores:
- your login details
- your city
- taxi phone numbers you added
- friends you shared info from the app with
- locations you checked for a taxi
- your reviews about taxis
Developer stores:
- GPS coordinates of all routes you travelled
Amazon cloud
Developer stores:
- your login details
- your city
- energy usage details
dedicated server park

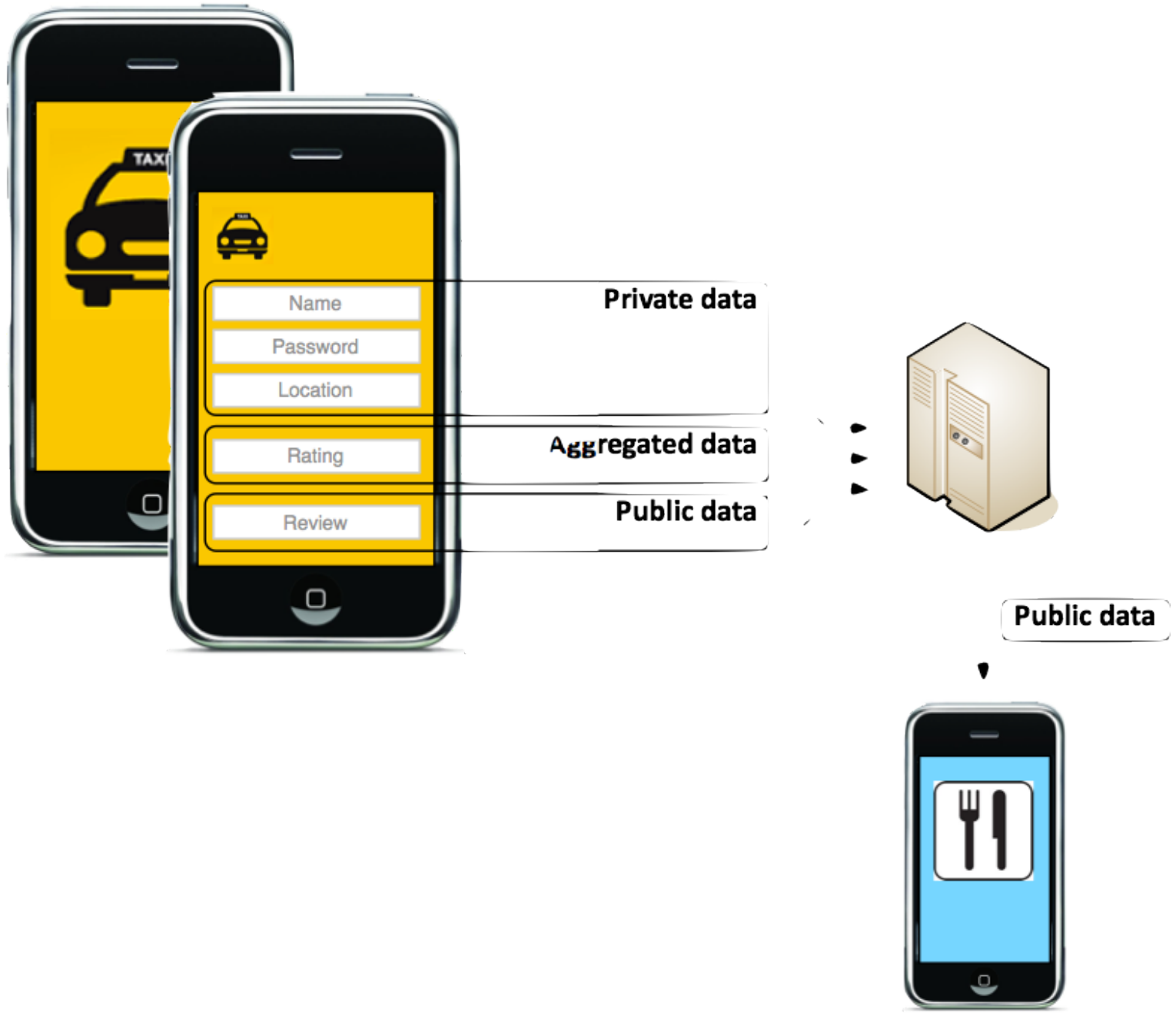
Wat is het

- De SDS is een (web)service voor persoonlijke opslag van gegevens die het mogelijk maakt delen van persoonlijke data openbaar te maken en andere delen privé te houden. In een toekomstige versie breiden we deze twee toegangsniveaus uit, zodat er meer gedetailleerd controle mogelijk is.
- Ontwikkeling van de SDS is geïnspireerd op een aantal verschillende concepten zoals: digitale kluis, vendor relationship management ideeën, crowd-source data verzamelen, citizens science projecten en de open data discussie in het algemeen.
- De SDS is een eerste stap in het combineren van open, publieke gegevens (niet noodzakelijk alleen de overheidsgegevens) met persoonlijk en privé (of zakelijke) gegevens. Door middel van het aanbieden van een API wordt het makkelijker voor ontwikkelaars slimme applicaties te ontwikkelen waar zowel private als publieke gegevens worden gecombineerd; met voordelen voor zowel de gebruiker als de ontwikkelaar en het publiek in het algemeen.


Taxi App
your data
Train Delay App
your data
Bike Rental App
your data
shared data store
trusted party

Taxi App

- Een taxi kan gebeld worden, waarbij de keuze gebaseerd is op beschikbare ratings
- Gebruikers geven automatisch informatie door; waar rijdt een taxi, beoordeling van de chauffeur, rating van de rit, prijs etc.
- De Taxi app houdt gegevens over de gemaakte ritten bij zoals datum, vertrek, bestemming etc, voor de persoonlijke doeleinden of bijvoorbeeld facturering.
- Bovenstaand voorbeeld laat zien dat een eenvoudige app data genereert die zowel interessant is voor een breed publiek als data die een gebruiker niet noodzakelijk wil delen. De SDS geeft ontwikkelaars de mogelijkheid deze scheiding van gegevens aan hun gebruikers aan te bieden.

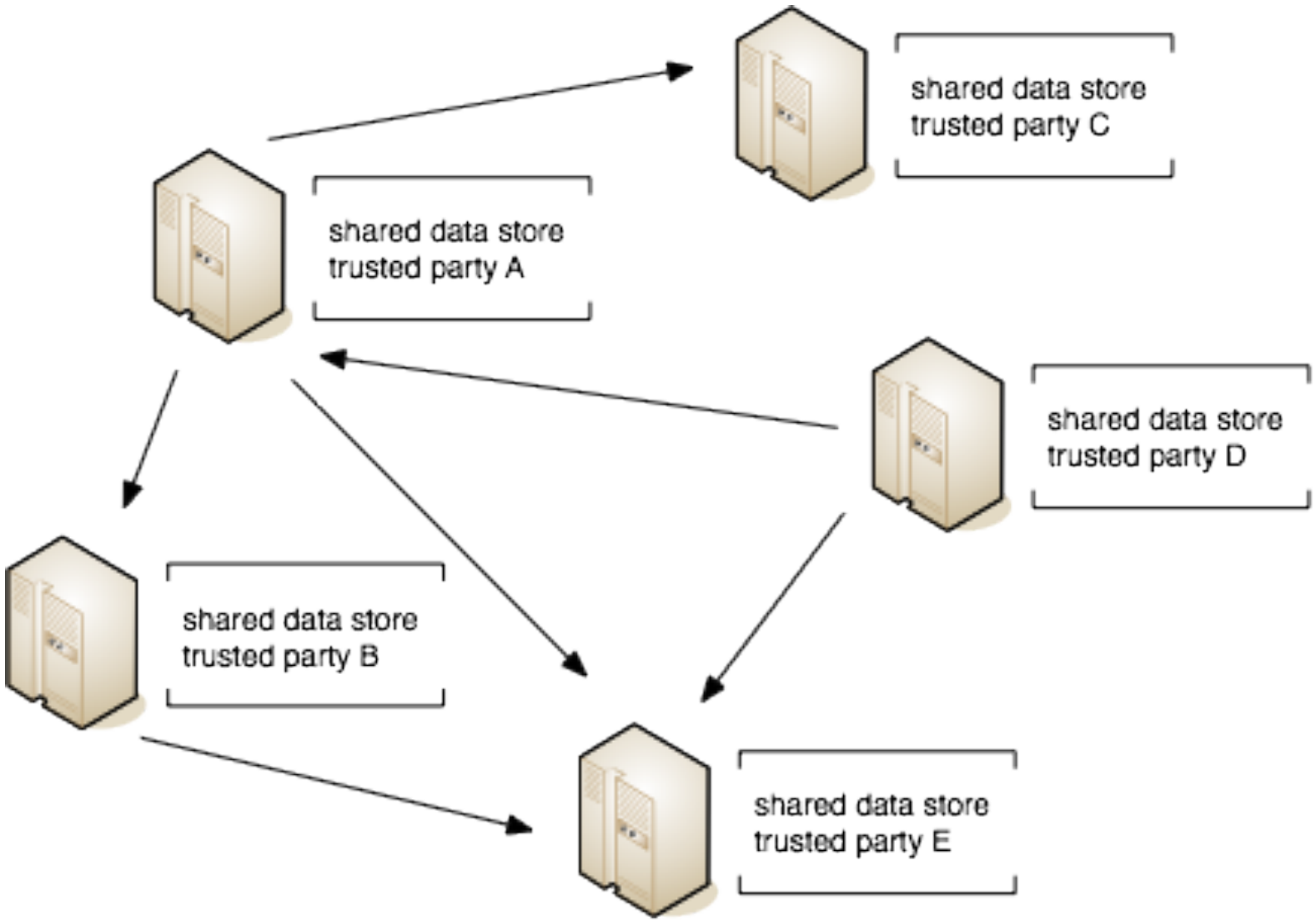

Name
Password
Location
Rating
Review
Private data
Aggregated data
Public data
Public data

Hoe het werkt

- De SDS is geïmplementeerd op een CouchDB database. Alles wordt opgeslagen per persoon, per context. Niet-openbare documenten in encrypted vorm: AES-256. Beheer van keys is zo ingericht, dat ze nooit op de server opgeslagen worden in een niet gecodeerde vorm. Een document wordt opgeslagen samen met een decoding key die is encrypted met de publieke key van potentiële lezers (classes van lezers).

Hoe het werkt

- Dit houdt in dat niemand de gegevens can decoderen behalve de beoogde partijen, ook de beheerder van de server niet. De data is veilig zelfs als de server problemen ondervindt. Als een key achterhaald word, bijv. met behulp van brute kracht of andere middelen, zijn de gegevens die daarmee gevonden worden beperkt tot een enkel document; er is geen master key.
- Dit betekent ook dat de verantwoordelijkheid voor het beheer van de keys van een individuele gebruiker bij deze gebruiker ligt. Het verlies van deze keys betekent het verlies van de gegevens; er zijn geen achteringangen. Applicatie-ontwikkelaars moeten zich bewust zijn van deze kwestie en de hier in de gebruikerservaring rekening mee houden (zonder in te boeten veiligheid). Extra diensten kunnen mogelijk aan de SDS worden toegevoegd om ontwikkelaars hierbij te helpen.


shared data store
trusted party C
shared data store
trusted party A
shared data store
trusted party D
shared data store
trusted party B
shared data store
trusted party E

Datamodel

- Het datamodel is free form, zoals gebruikelijk is bij document stores. De SDS heeft uiteraard wel een structuur. De gegevens worden opgeslagen in een 'context'. Elke context kan een bepaalde structuur afdwingen, door de context eigenaar gespecificeerd als een javascript validatie functie. Dit zou een 'minimale' structuur zijn. Extra velden zijn altijd mogelijk en hebben geen validatie nodig. In het uiterste geval is de validatiefunctie leeg. In dat geval is de aanwezigheid van bepaalde velden niet gegarandeerd en applicaties die voor deze context worden geschreven dienen hierop aangepast te worden.

Datamodel

- Elke context heeft ook een 'publish' functie. Default worden alle gegevens privé; depublish JavaScript-functie haalt gegevens eruit die publiek gebruikt mogen worden. In de toekomst kunnen de gegevens voor specifieke toepassingen van bijvoorbeeld de ontwikkelaar gebruikt worden, door middel van de PP key pair van de applicatie. Elke context kan door meerdere applicaties gebruikt worden. Deze apps hebben hun eigen PP key pair en delen de data omgeving van de context. Aangezien de dat free form is, kunnen ze nog steeds specifieke functionaliteit bieden. Ze delen ze de validatie en publish functies van de context.