Panorama du Big Data

Pourquoi maintenant ?

- coûts de stockage

- puissance de calcul

- le débit

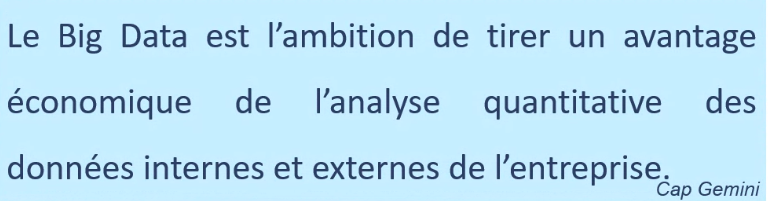
Datalake : stocke les données à l’état brut

Les 3 V :

- volume

- vélocité

- variété



ETL : extract transform load

W3C : organisme pour les guides et protocoles du web

Etude de cas :

Identifier les entreprise qui vendent.

Données internes :

- demander l’identité bancaire personnelle de locataire en plus de celle de son entreprise (si c’est légal) et la conserver afin de pouvoir lui facturer les dépenses en électricité des derniers mois s’il décide de partir sans payer ce qu’il reste.

- conserver le numéro de téléphone et l’adresse mail personnels du locataire en plus de son identité bancaire afin de pouvoir le conatcter si besoin.

Données externes :

- obtenir l’identité du nouveau locataire en faisant appel à un prestataire : l’agence qui loue le local.