Plan du cours :

Introduction :

Qu’est-ce que le big data ?

Commentaires laissés sur les réseaux sociaux ? Facebook ou Twitter, vidéos publiées sur YouTube ou Dailymotion, musique sur iTunes ou Deezer, paiements bancaires, emails, signaux GPS, carte SIM des appareils électroniques... : plus de **2,5 trillions d'octets de données transitent sur le web chaque jour**, notamment grâce à la démocratisation du digital depuis l'ADSL et l'apparition des smartphones et tablettes

e **Big Data** se caractérise par la problématique des 3V :

* **Vélocité**: la vitesse à laquelle les données sont traitées simultanément
* **Variété**: l'origine variée des sources de données qui arrivent non structurées (formats, codes, langages différents...)
* **Volume**: le poids total des données collectées

Ces 3V soulèvent la question de la gestion et de l'**exploitation de données**. Comment gérer des données de plus en plus nombreuses, de plus en plus diverses, et le faire de plus en plus rapidement ? Il faut en effet penser à la façon de les collecter, de les stocker puis de les analyser pour les exploiter à bon escient.