

Introduction to Data Science

Qu'est-ce que la Data Science ?

Définition

Ensemble d'outils et de techniques et qui permettent de **collecter, nettoyer, transformer, analyser et tirer parti des données** afin d'en dégager des enseignements.

Compétences requises

- Mathématiques (statistiques, probabilités, modélisation)
- Programmation informatique (Python ou R)
- Réseau (Architecture Cloud, API, etc)

Introduction to Data Science

Les facteurs de l'essor de la Data Science

- Explosion de la génération des données sur le web (GAFAM, BATX..) et grâce aux objets connectés
- Grande capacité de stockage des données grâce à l'écosystème Big Data
- Augmentation considérable de la puissance de calcul
- Baisse exponentielle des coûts de stockage et de calcul

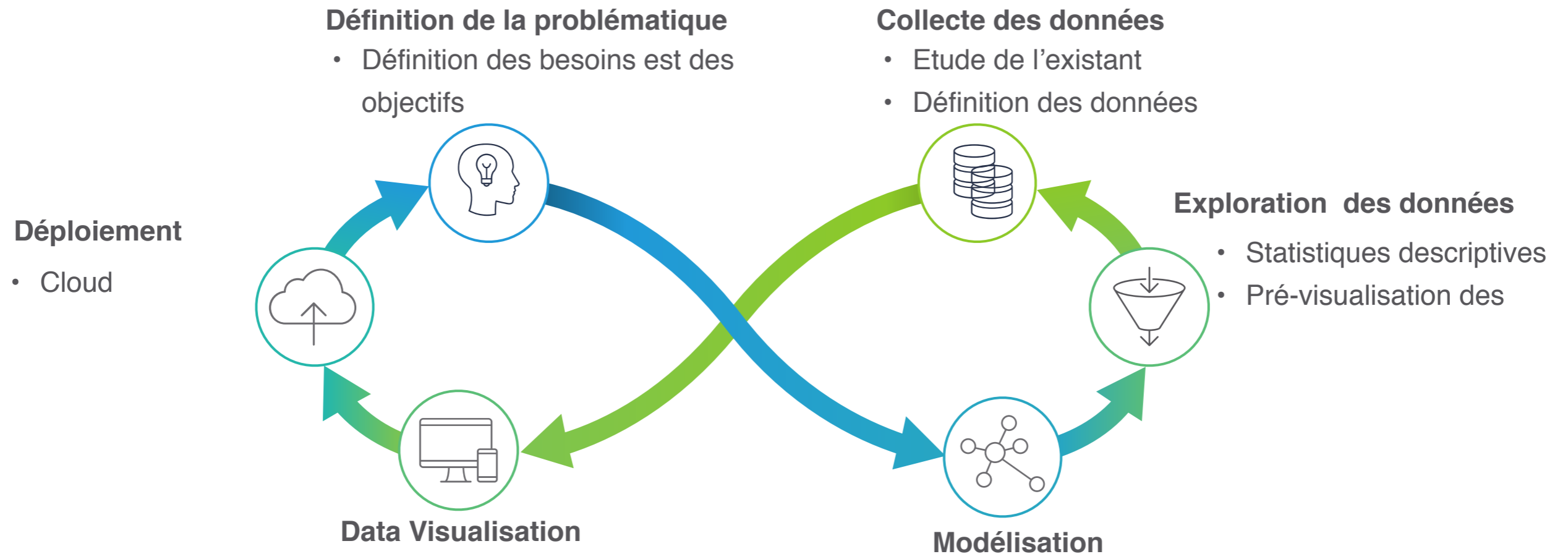
Introduction to Data Science

Les métiers de la Data Science

- Data Scientist
- Data Analyst
- Data Engineer
- DevOps
- Métier

Introduction to Data Science

Le cycle de vie d'un projet en Data Science



Introduction to Data Science

Focus sur la gouvernance et la qualité des données

- Définition des politiques de gouvernances des données:
- Rôles et Responsabilités
- Processus métiers et données associées
- Définition et implémentation des contrôles de la qualité des données
- Revue périodique et/ ou automatique de la qualité des données
- Rôle de Chief Data Quality Officer