

Le *big data* en contexte de  
transformation numérique:  
facteur de succès ou facteur de risque?

Grégory Vial  
gregory.vial@hec.ca

*“Big data is like teenage sex: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it...”*

(Dan Ariely)

# Présentation

- Professeur adjoint, HEC Montréal
- Principaux intérêts de recherche:
  - Développement de systèmes (ex. agile, lean, DevOps)
  - Gestion des données (transactionnel / analytique)
- 8 ans de carrière dans le développement logiciel (logistique)
  - Spécialisation en gestion des données

# Transformation numérique?

- Un phénomène nouveau?
  - *Quo vadis* Michael Hammer, Venkatraman et consorts?
- Évolution du phénomène des années 90:
  - Étendue plus vaste (processus → modèle d'affaires)
  - Origine différente (décision organisationnelle → tendance globale)
  - Types de technologies (ex. PGI → *combinaisons* de technologies)
  - Incertitude (interne → externe)
  - Défis (locaux → globaux)

# Transformation numérique?

## Caractéristiques

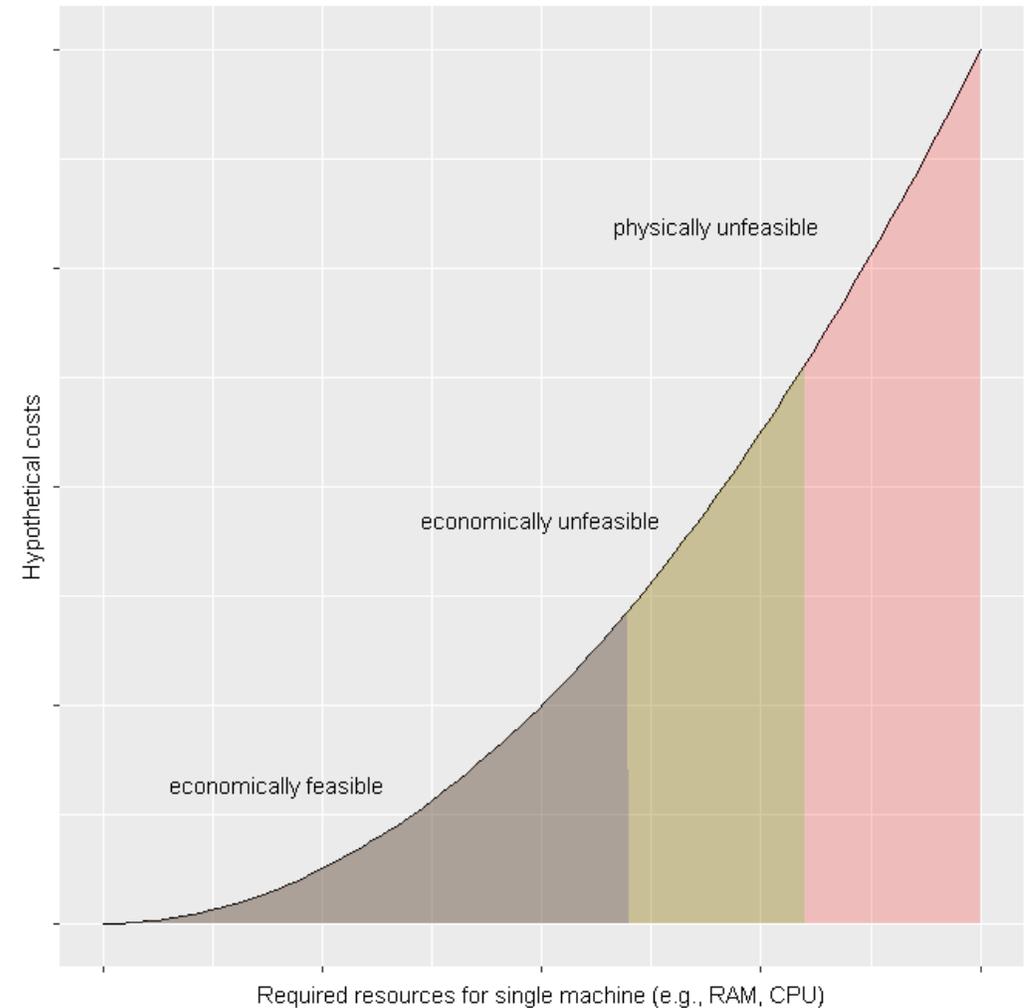
- Processus
- Efforts à grande échelle
- Changements fondamentaux
- Perte de contrôle (relative!)
- Bénéfices à horizons multiples
- Technologies multiples
- Évolution accélérée (besoins, technologies)
- Impacts à différents niveaux

## Implications pour la gestion de projet

- Gestion par projet, produit ou autre?
- Arrimage avec les parties prenantes
- Résistance, inertie organisationnelle
- Gestion des risques
- Justification sur différentes échelles
- Multidisciplinarité, intégration
- Maintien d'un équilibre sain entre flexibilité et engagement
- (Ré)Conciliation des bénéfices

# Big Data?

- 4 grandes caractéristiques:
  - Volume
  - Vitesse
  - Variété
  - Véracité
- L'idée est de créer de la valeur à travers l'exploitation des données



# Le big data, pour quoi faire?

- 3 grands types d'analyses:

1. **Descriptive**: Dresser un portrait de la situation actuelle/passée

2. **Prédictive**: Utiliser les données générées dans le passé pour prédire des événements

3. **Prescriptive**: Générer des recommandations personnalisées pour guider la prise d'action, parfois sans intervention humaine

# Le big data, un passage obligé?

- Pas nécessairement mais la transformation numérique peut:
  - Être appuyée par les données (antécédent)
  - Générer un impact sur la génération de données (conséquence)
  - Ou les deux en même temps!
- Les technologies utilisées dans nos sociétés favorisent la génération de données
  - Appareils mobiles, internet des objets
- Elles favorisent également l'accès et le partage des données
  - Plateformes et écosystèmes numériques, médias sociaux

# Big data, entre mythes et réalité

- Chez les organisations qui réussissent:
  - Le big data au cœur des compétences
  - Des moyens financiers importants (souvent)
  - Un impact sur des ressources matérielles (ex. industrie 4.0)
  - Une stratégie d'affaires viables qui exploite les données (et non l'inverse!)
- Et chez les autres:
  - Gartner rapporte environ 60% de taux d'échec (jusqu'à 85%)

# Cas: Port de Montréal

- Objectifs

- Augmenter l'efficacité et l'efficience du flux des marchandises
- Réduire les émissions de gaz à effet de serre
- Développer une compétence exportable à l'étranger

- Défis

- Connexion d'équipements existants
- Reconnaissance du potentiel de valeur des données
- Amélioration des processus dans un système complexe et réglementé

# Cas: Transport for London

- Objectifs

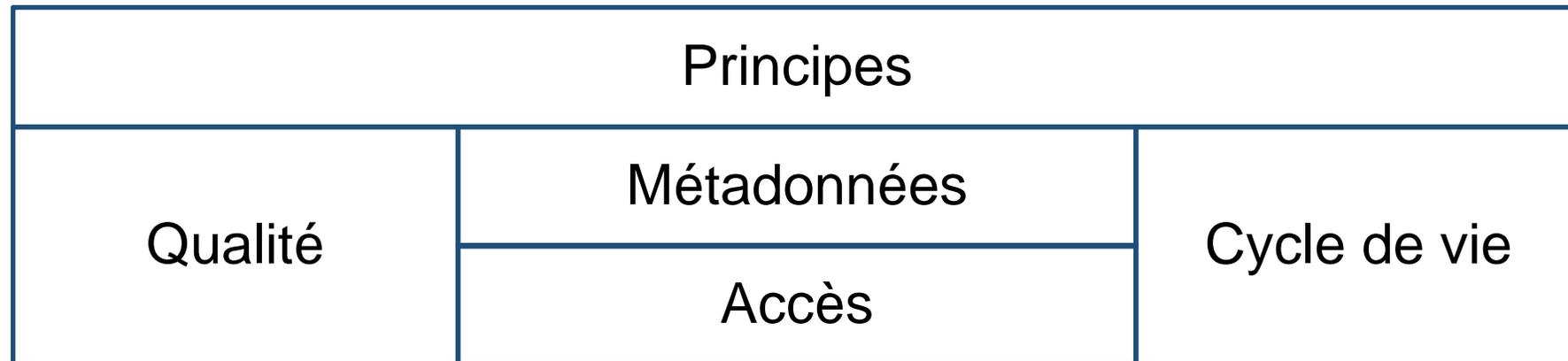
- Comprendre comment les gens utilisent le transport en commun
- Redistribuer le trafic en temps réel
- Aider la planification urbaine

- Défis

- Processus parfois « déconnectés » (bus versus métro)
- Inférence à partir de données manquantes
- Assurer la pérennité de l'initiative à travers différentes administrations

# Gestion des données

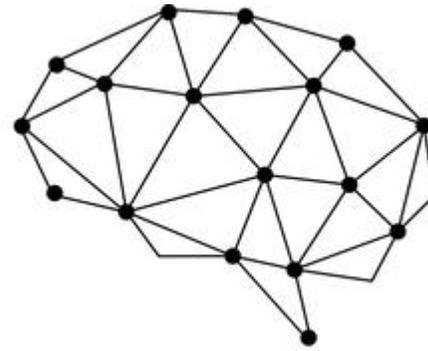
- Définition d'un cadre de gouvernance:
  - Qui possède l'autorité pour prendre des décisions?
  - Qui est responsable une fois ces décisions prises?



D'après Khatri & Brown (2010)

# Gouvernance des données

- Principes
  - ex. utilisation, reconnaissance comme actif stratégique, cadres légaux?
- Qualité
  - ex. standards, précision, complétude, crédibilité?
- Métadonnées
  - ex. sémantique, documentation, mise à jour?
- Accès
  - ex. gestion des risques, conformité, audit, restauration?
- Cycle de vie
  - ex. production, rétention, purge, relation avec des impératifs légaux?



Cambridge  
Analytica

*“Chaque jour, nous disséminons des données ultra-personnelles ; chaque jour, ces données sont stockées, utilisées, croisées au profit de quelqu'un, quelque part. Et ce quelqu'un, ce n'est pas vous.”*

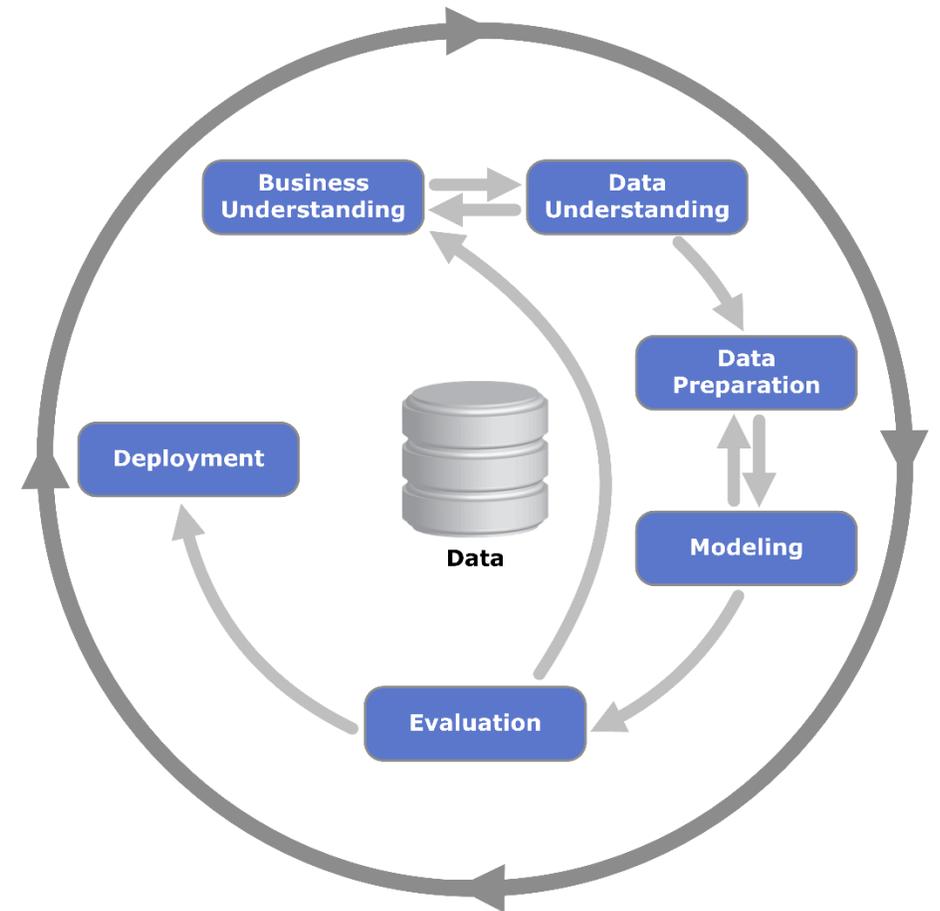
(P. Lagacé, La Presse, 07/03/2018)

# Enjeux éthiques et légaux

- Peu considérés jusqu'au scandale de CA
- Doivent être considérés comme éléments à part entière d'un projet
  - Gestion des parties prenantes
  - Gestion des facteurs de risque
- Quelques exemples:
  - Politique d'accès
  - Partage direct des données (ex. Amazon S3, Google Drive)
  - Partage des données par API
  - Croisement de données

# Gestion de projet

- CRISP-DM (Cross-Industry Standard Process for Data Mining)
- Méthodologie la plus fréquemment utilisée pour les projets d'analyse de données
- Bien qu'intéressante, elle reste à un niveau élevé



# Tendances actuelles en gestion de projet

- Approches itératives et agiles afin de favoriser:
  - L'expérimentation
  - La multidisciplinarité et la collaboration
  - La création de valeur « rapide »
- Principaux défis:
  - Technologies parfois peu matures
  - « Stacks » technologiques complexes
  - Engagements (et investissements) sur des hypothèses



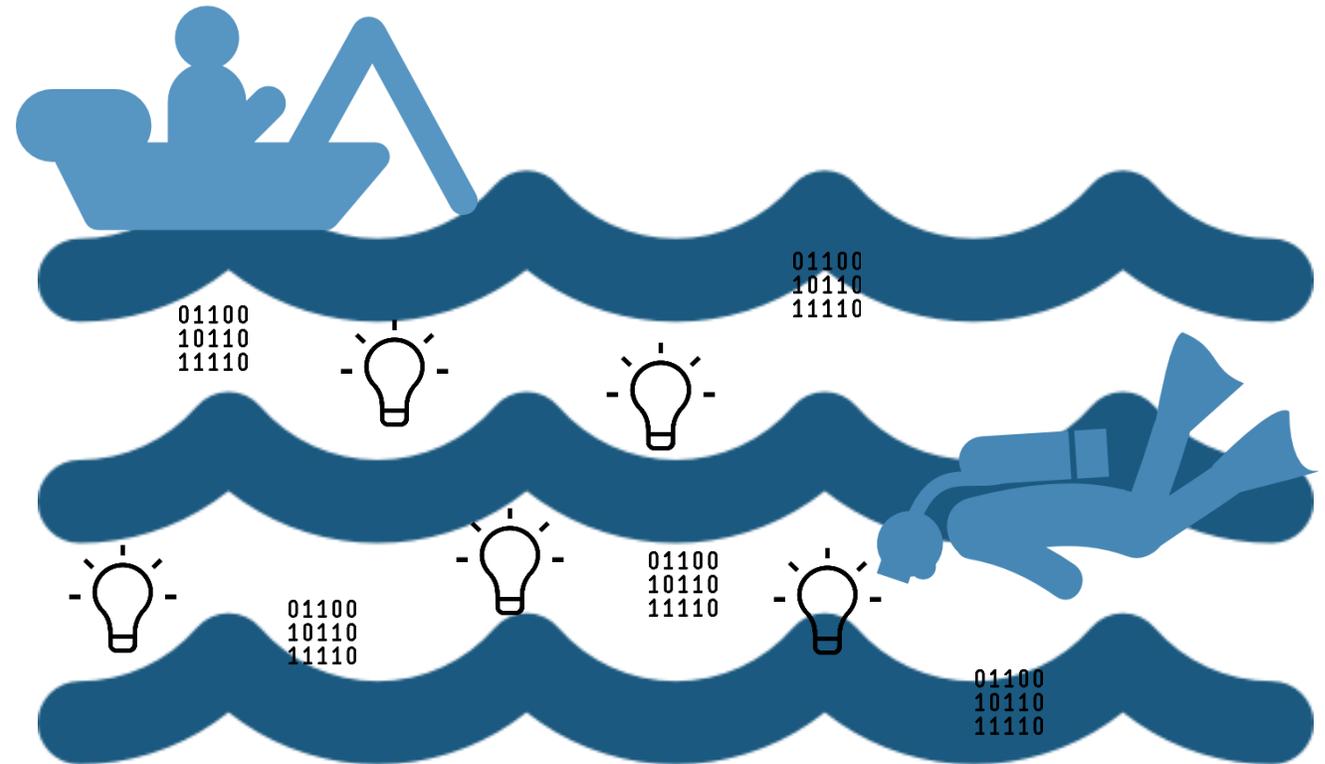
icône: Freepik ([www.flaticon.com](http://www.flaticon.com))

# Tendances actuelles en gestion de projet

- **Projet traditionnel:**
  - Analyse des besoins
  - Création d'une structure de données
  - Implantation du processus pour alimenter ces structures (ETL)
  - Exploitation des données
- **Tendance actuelle:**
  - Ingestion des données brutes
  - Transformation et exploitation des données au besoin (ELT)

# Tendances actuelles en gestion de données

- Un *data lake* est un espace où l'on regroupe des données brutes (ou presque)
- Ces données sont exploitées pour générer de la connaissance
- L'idée a été lancée en 2010 par le CTO de Pentaho



# Un data lake, oui mais...

*“Ce n’est pas parce qu’on vit près du lac qu’on mange du poisson quand on veut.”*

(proverbe africain)

Certains auteurs (ex. Bill Inmon) mettent en garde contre le danger de créer des dépotoirs de données



icône: Freepik (www.flaticon.com)

# Tendances actuelles en gestion de projet

- Avantages:

- Plus d'agilité
- Plus de possibilités
- La transformation est mise entre les mains des experts
- Coûts initiaux plus faibles (en théorie)

- À surveiller:

- Inefficiences
- Répétition
- Réservé aux experts

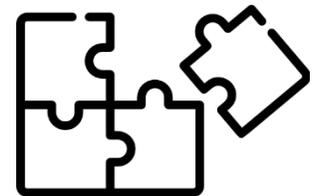
# Et le gestionnaire de projet dans tout ça?

- 2 philosophies (et ce n'est pas nouveau!)
  - Le spécialiste
  - Le généraliste
- Chaque approche possède ses avantages
  - Le spécialiste possède une connaissance technique précieuse
  - Le généraliste ne se perd pas dans la technologie
- ... et ses inconvénients
  - Le spécialiste pourrait moins comprendre la réalité de l'entreprise
  - Le généraliste pourrait se perdre dans le jargon technique complexe



# Les plus gros défis

- La technologie n'est qu'une pièce du puzzle
- Développer une culture de la donnée représente un changement organisationnel
  - Au niveau des pratiques et processus
  - Au niveau des mentalités
- Certaines organisations développent un leadership spécifique sur ce thème
  - Chief Digital Officer, Chief Data Officer



icône: Freepik (www.flaticon.com)

# Les plus gros défis

- Pour beaucoup d'organisations, le but est d'arriver à être ambidextre pour:
  - Exploiter les ressources actuelles (efficience, efficacité)
  - Explorer des nouvelles pistes (innovation)
- Difficile d'atteindre ces deux objectifs en même temps
- Le big data peut y contribuer mais...



Icône: Freepik (www.flaticon.com)

# Maturité numérique?

*“organizations don’t wave a magic wand and instantly become digitally adept. Instead, it’s better expressed as ‘maturity’. ”*

(Gerald C. Kane)

Maturité: Capacité à réagir d’une façon appropriée qui est généralement acquise et non innée



# En conclusion

- *Big* ou non, un défi de la transformation numérique est de développer une compétence pour exploiter les données
- Ceci représente dans beaucoup de cas un virage important
  - Passage de la génération de données à leur exploitation
- La littérature et les études de cas montrent que des approches itératives et évolutives ont plus de chances de succès

# Merci!

Gregory Vial (gregory.vial@hec.ca)