



ETAT DE L'ART DE LA VALORISATION DES DONNEES

Café Atelier PIGMA – Jeudi 10 Octobre 2019

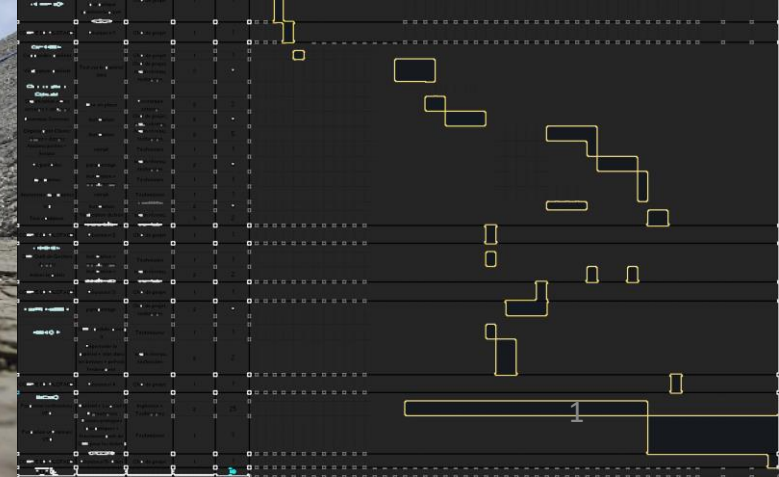
Dr. Antoine Jeanjean

Bordeaux Data Science / Recommerce

antoinejeanjan@gmail.com



Diffusion sans accord interdite





- La donnée est aujourd'hui **le nouvel or noir** des sociétés, des organisations, des administrations et des états.
- Comme toute ruée vers l'or, on assiste à une « **montée en pression** » : *hystérie médiatique, formations légères, consulting carton-plâtre, outillage fumeux ou trop complexe, projets pharaoniques démesurés, précipitation...*
- **Mes 1^{er} conseils :**
 - *Gardez votre calme, respirez et prenez votre temps !*
 - *Formez-vous, lisez différents points de vue, documentez-vous ;*
 - *Ne prenez pas de décisions trop rapidement sur ces sujets cruciaux ;*
 - *Ne restez pas immobiles par peur du changement sur ces thématiques complexes ;*
 - *Définissez bien les objectifs de votre projet de DATA au préalable.*



Maîtrisez vos données



- Une bonne maîtrise des flux de la donnée va vous permettre :
 - De **connaître** l'information : quels types de données ?
 - De **localiser** l'information : où sont vos données ?
 - De **sécuriser** l'information : vous maîtrisez qui voit quoi / vous êtes conforme à RGPD
 - De **mettre à disposition** l'information : vos données ne dorment pas
 - **D'utiliser** l'information : vos données servent à prendre des décisions
 - De **calculer** de nouvelles informations : vous fabriquez de prévisions / indicateurs
 - De prendre au final de bonnes décisions en **UTILISANT** vos données

→ Gestion du « stock » de données



- La donnée n'est pas encore **valorisée** (en général) dans le bilan d'une entreprise où figure pourtant le stock de marchandises ;
- On peut s'attendre à ce que le **STOCK** de données soit un jour (prochain) valorisé comptablement par les directions financières (*quand elles auront compris les enjeux et le potentiel sous-jacent !*) ;
- Comme votre stock physique :
 - Si on vous vole des données → **risque** pour l'activité ;
 - Si vos données sont de mauvaise qualité → **risque** pour l'activité ;
 - Si vos données sont périmées → **risque** pour l'activité ;
 - Vous allez pouvoir **valoriser / commercialiser** vos données
→ des sociétés spécialisées (marketplace de données) voient le jour pour que vous puissiez vendre vos données ;
 - Vous allez devoir apprendre à **ranger / structurer / organiser / mettre à disposition et bonifier** vos données
→ maximiser les valorisations de ce stock dormant
- On voit même à des sociétés qui commercialisent à perte un service afin de revendre ensuite les données associées au service qu'ils vous ont mis à disposition → la valorisation est **intégrée** au business model (Ex : Google)

→ Cartographie des disciplines DATA



Source : @toniojj

→ Positionnement de l'analyse de données



1. Phase de ***récupération des données*** : internes (tracker, S.I. métiers, scripts d'automatisation...) ou externes (scrapping, achat de données, ...) ;
2. Phase de ***structuration des données*** : nettoyage des données (redondance, données erronées), indexation, matching, optimisation du stockage, sécurisation, cartographie ;
3. Phase de ***mise à disposition des données*** : sous forme de tableaux de bords ou de rapports de Business Intelligence.

Ingénieurs Informatiques
Développeurs

Data Engineer
Administrateur
de base de données

Data Engineer
Consultant BI



4. Phase de **création d'analyses** pour créer des indicateurs ou des prévisions, grâce aux Statistiques :

- **Modélisation des données** : Régressions, ajustement de loi de proba, analyse de variance, ...
- **Analyse de donnée** : Analyse en composantes principales, analyse factorielle, analyse des correspondances multiples, Classification, partitionnement, ...
- **Tests** : de corrélation, de tendance, paramétriques, non paramétriques, ...

Data Analyst
Statisticien
Data Scientist

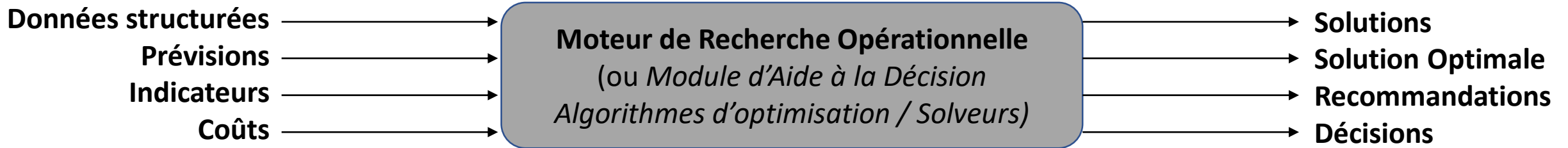
→ Les **statistiques** sont de formidables outils permettant de calculer des RICH DATA utilisées ensuite par des outils d'aide à la décision.



5. La Recherche Opérationnelle agit ensuite et va se nourrir de ces rich data.

Research Engineer
Revenue Manager
Ingénieur informatique

La Recherche Opérationnelle est une discipline qui s'appuie sur des méthodes mathématiques et qui permet d'aider à prendre de meilleures décisions grâce à la modélisation puis la résolution d'un problème d'optimisation qui tend à maximiser un objectif tout en respectant un ensemble de contraintes.



Intelligence Artificielle
"Remplacer l'humain"

≠

Recherche Opérationnelle
"Seconder l'humain"

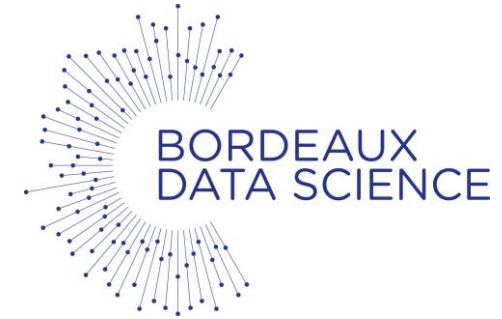


La Data Science à Bordeaux



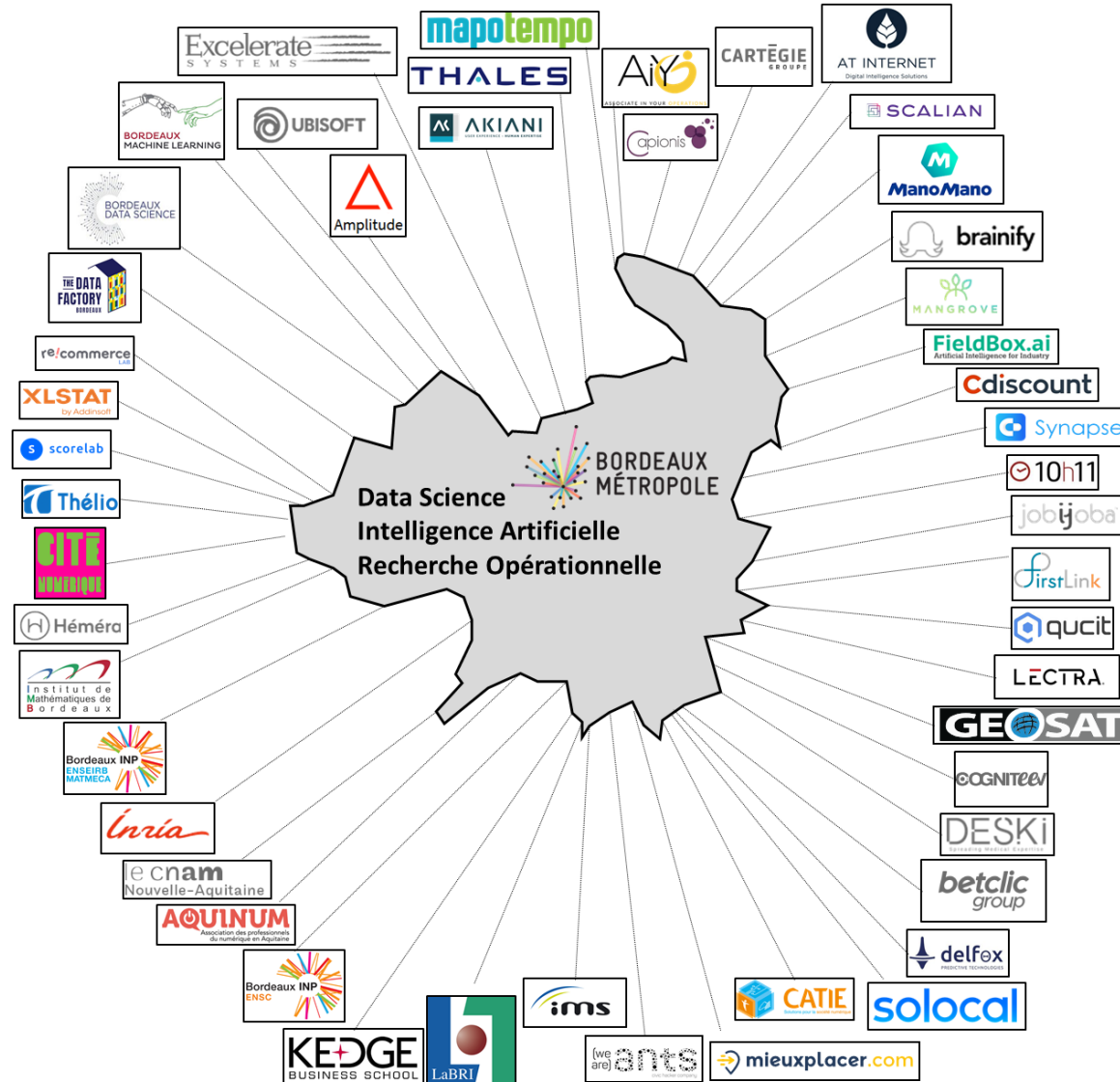
- **Bordeaux Data Science :**

- Plusieurs sociétés membres :
XLStat, Recommerce, Scorelab, Synapse, Thelio, Cdiscount, Avisia...
- Plusieurs écoles impliquées :
Université Mathématiques / Statistiques de Bordeaux, Kedge Business School,
- 1 meetup tous les 15 jours à la **DATA FACTORY** – 35 rue Thiac Bordeaux
- **Objectifs :**
 - Partager sur les thématiques de Statistiques / Recherche Opérationnelle / I.A / ... ;
 - Présenter de nouveaux problèmes ;
 - Echanger sur les techniques, les modélisations ou les méthodes de résolution ;
 - Présenter les travaux de thèse d'étudiants ;
 - Découvrir les sociétés de conseil du secteur ;




www.bordeauxdatascience.fr

Un éco-système dynamique



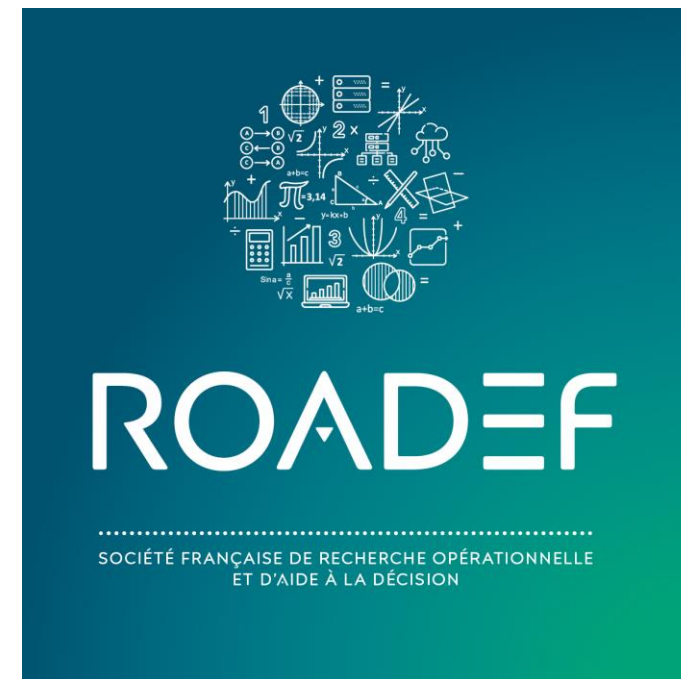
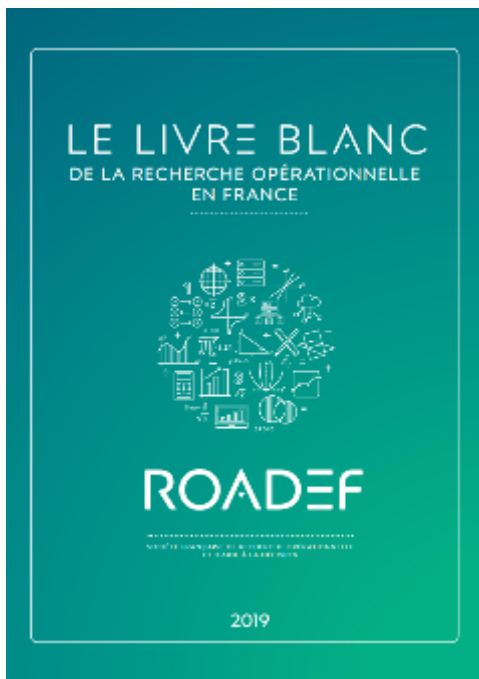


- Société Française de Recherche Opérationnelle : 
- 1000 membres ;
- 1 association avec un bureau de 7 membres ;
- 1 congrès annuel avec 600 participants ;
- Plus de 40 sociétés membres dont **ENGIE, EDF, Air France, Google, SNCF, Veolia, Bouygues, Schneider, Amadeus, Renault, RATP, ...** mais aussi des **PME**
- Toutes les écoles d'ingénieurs, de commerce et les universités représentées ;
- L'objectif :
 - Partager sur les thématiques de Recherche Opérationnelle ;
 - Présenter de nouveaux problèmes ;
 - Echanger sur les modélisations ou les méthodes de résolution ;
 - Présenter les travaux de thèse des nouveaux étudiants ;
 - Première place de l'emploi en terme de RO ;
 - Découvrir les sociétés de conseil du secteur et les logiciels solveurs.

www.roadef.org



La R.O. en France



Sortie du Livre Blanc de la Recherche Opérationnelle
en France fin Mai 2019

Disponible gratuitement sur www.roadef.org



Dr. Antoine Jeanjean
Association Bordeaux Data Science
Recommerce Group
antoinejeanjean@gmail.com

www.bordeauxdatascience.fr



ANNEXES

Exemple d'un projet DATA

