

Animation de data scientists en entreprise : comment ça marche ?

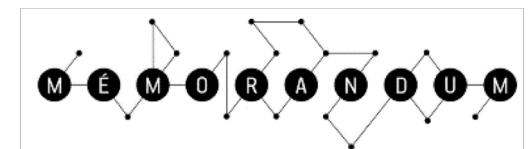
Denis Oblin

Denis.oblin@memorandum.pro

Romain Jouin

romain.jouin@memorandum.pro / 0652868730

mars 2016



Comment animer des équipes de datascientists ?

Qui se pose la question ?

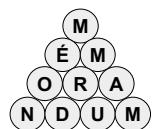
De grandes entreprises avec des équipes éclatées ?

Des entreprises qui ont un ou deux spécimens ?

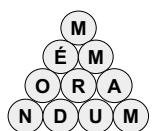
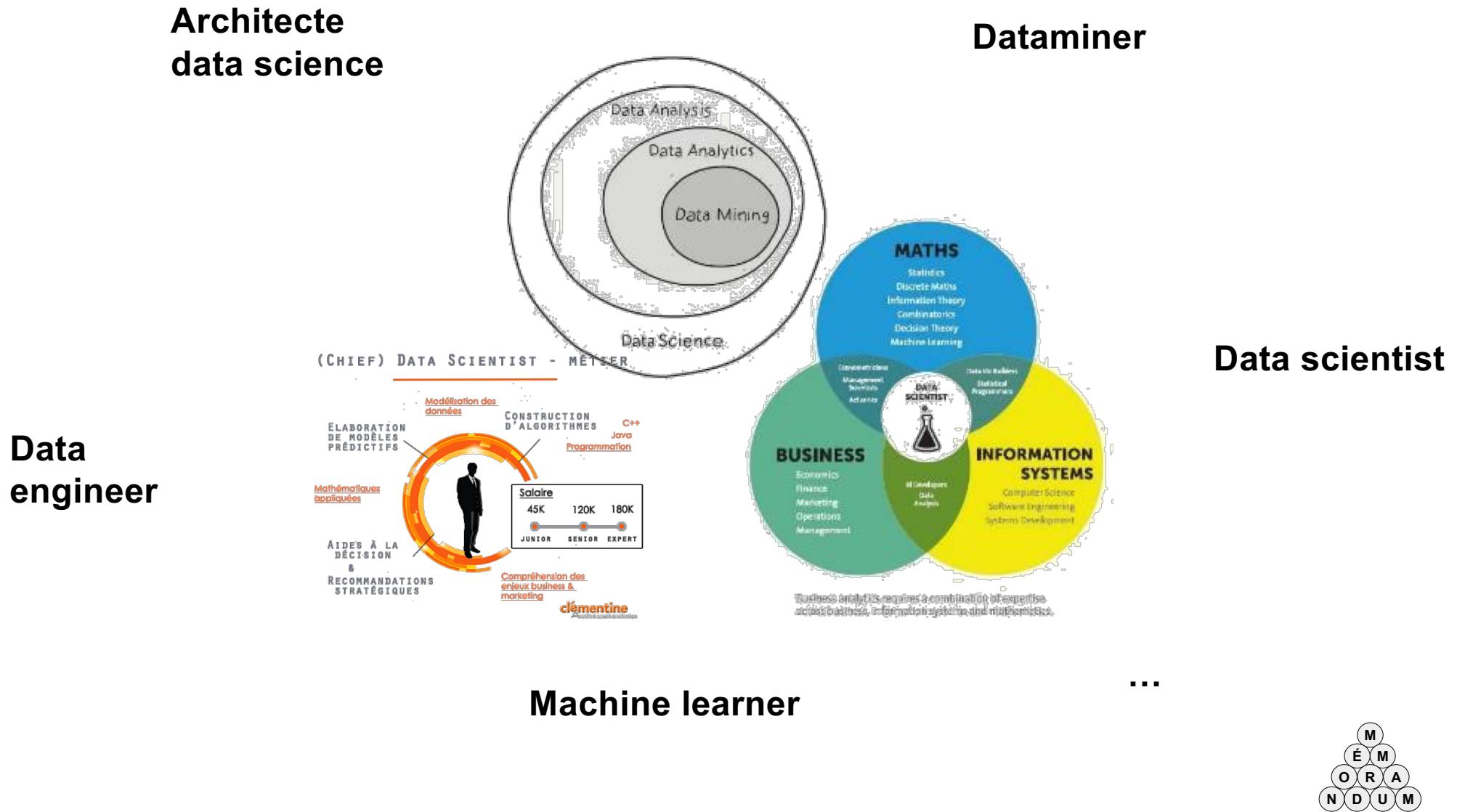
Des entreprises qui veulent faire émerger des talents ?

Qui ne se la pose pas ?

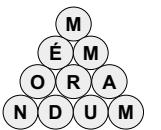
Des entreprises qui achètent de la magie



De qui parle-t-on ?



Management des datascientist : pour quoi faire ?



Plus précisément, quelle question de management ?

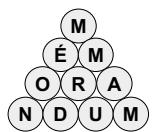
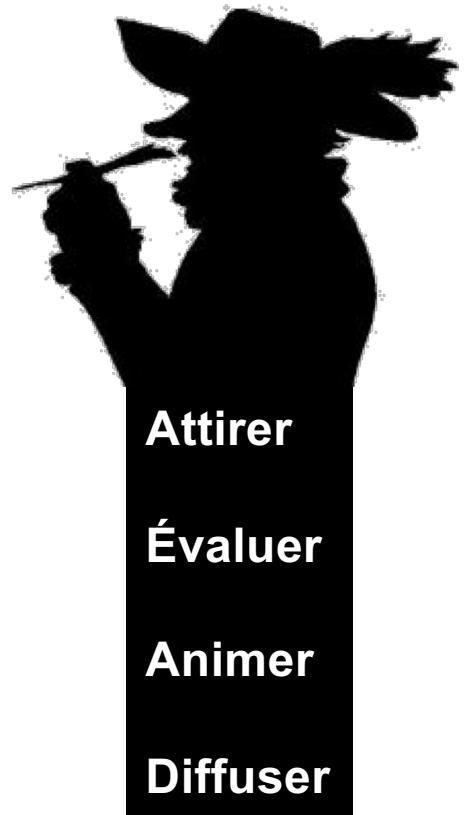
« Animer pour encourager » ? Ah ! Non ! C'est un peu court, jeune homme ! On pouvait dire... oh ! Dieu ! ... bien des choses en somme... En variant le ton, —par exemple, tenez

Coach : « mes talents vont s'étioler sans émulation ni soutien entre eux. Aucun ne connaît le champ global de la datascience : → tous seuls ils vont plus vite, ensemble ils iront plus loin »

Recruteur : « Comment convaincre un talent à nous rejoindre si nous ne lui montrons pas un environnement interne favorable à son épanouissement : »

Manager : « je ne comprends rien à ce que fait mon 'datascientist'. C'est sûrement magique mais comment challenger, backuper, évaluer ? »

Gestionnaire de risque : « il faut documenter les travaux de ces énergumènes sans quoi ils nous laisseront sur la paille dès qu'ils changeront de crèmerie »



Quels livrables pour une équipe data ?

Volet performance opérationnelle

A court terme la porte d'accès au machine learning est technique.

Une équipe dédiée au démarrage pour développer cette expertise

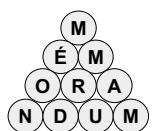
Enjeu (les vrais livrables)

- 1 Garantir l'autonomie technique (à la fois algorithmique et gestion spécifique de projet data)
- 2 Diffuser la culture : celle d'apprendre à poser les questions plus encore que celle d'apprendre à y répondre
- 3 Diagnostiquer l'accès aux données

Volet évolution stratégique

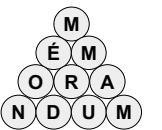
Développer de nouveaux services pour nos clients (exploitant la donnée)

Vendre nos données (pas les données perso mais des agrégats qui ont du sens pour des partenaires spécifiques)

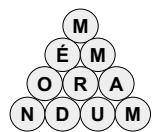


D'autres enjeux liés à l'animation de data scientists ?

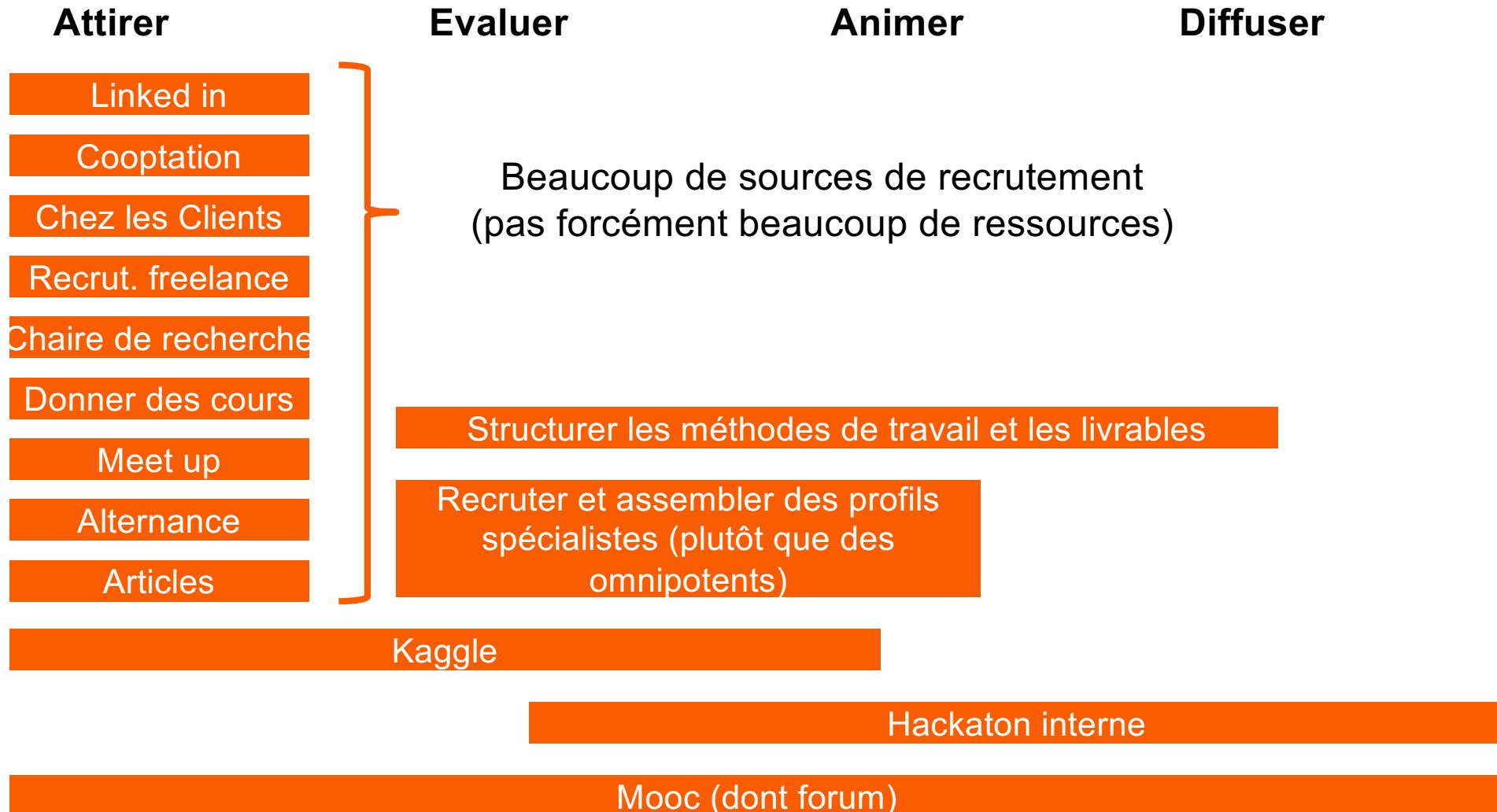
Séquence crowdsourcing



Quels leviers d'action ?



Des leviers d'action ?



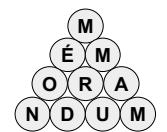
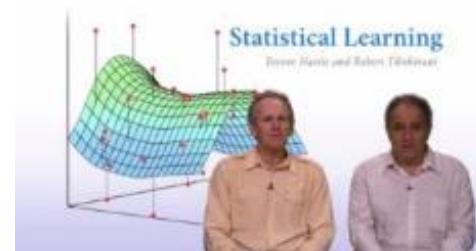
Exemple de leviers : Des moocs internes

Principe lancer une formation interne s'appuyant sur un mooc, ouvert à tous, pouvant être réalisé en partie sur le temps de travail

- Tout le monde apprend un peu
- Quelques uns vont au bout
- Certains sont bons et veulent continuer

L'entreprise

- a passé son message interne sur l'importance qu'elle accorde aux données
- Donne sa chance à tous
- Fait émerger un pool de salariés motivés .. Et reconnaissants



Exemple de leviers : Plateforme de concours

La participation de salariés aux concours en ligne réponds à beaucoup de besoins

- Alimenter le data scientist avec des sujets sympa
- Donner au management une échelle d'évaluation de son équipe
- Rendre la marque visible auprès d'une communauté (venez chez nous : on a des bons qui vous feront monter en compétence / vous pourrez continuer à briller ..)

Kaggle

The Home of Data Science

COMPETITIONS • CUSTOMER SOLUTIONS • JOBS BOARD

[Get started »](#)



Datascience.net (racheté par EY)

Mettez votre talent de data scientist au service des entreprises

Postulez à des challenges en ligne, trouvez le meilleur modèle, gagnez jusqu'à 10% de votre participation et devenez un expert.

Emmenez vos connaissances et compétences ?

[Je m'inscris](#)

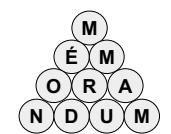
1. Téléchargez les données

2. Construez votre modèle

3. Testez la performance

Challenges terminés

Créez votre compétition



Exemple de leviers :

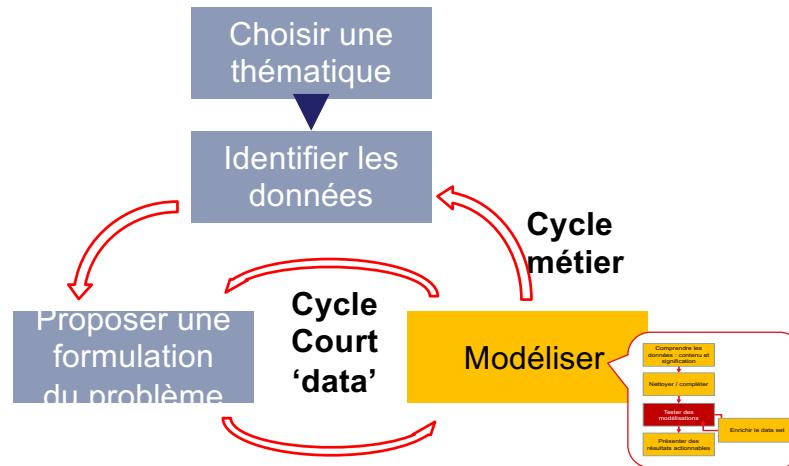
Structurer les méthodes de travail et les livrables

Dérouler une structure de travail adaptée des méthodes agiles / scrum

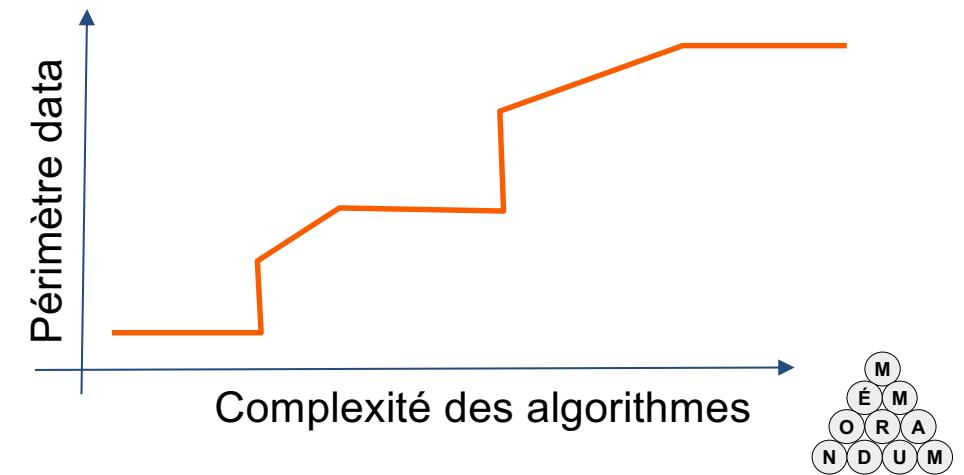
- Rend visible l'avancée des travaux

Adaptations possibles de la méthode agile

- Des users story itératives plutôt que successives
 - Faire progresser la réponse en terme de complexité algorithmique et périmètre data pris en compte



- Faire progresser la question en même temps que la réponse



Exemple de leviers :

Structurer les méthodes de travail et les livrables

Quelle trace un data scientist laisse t il de son travail ?

Les résultats opérationnels :

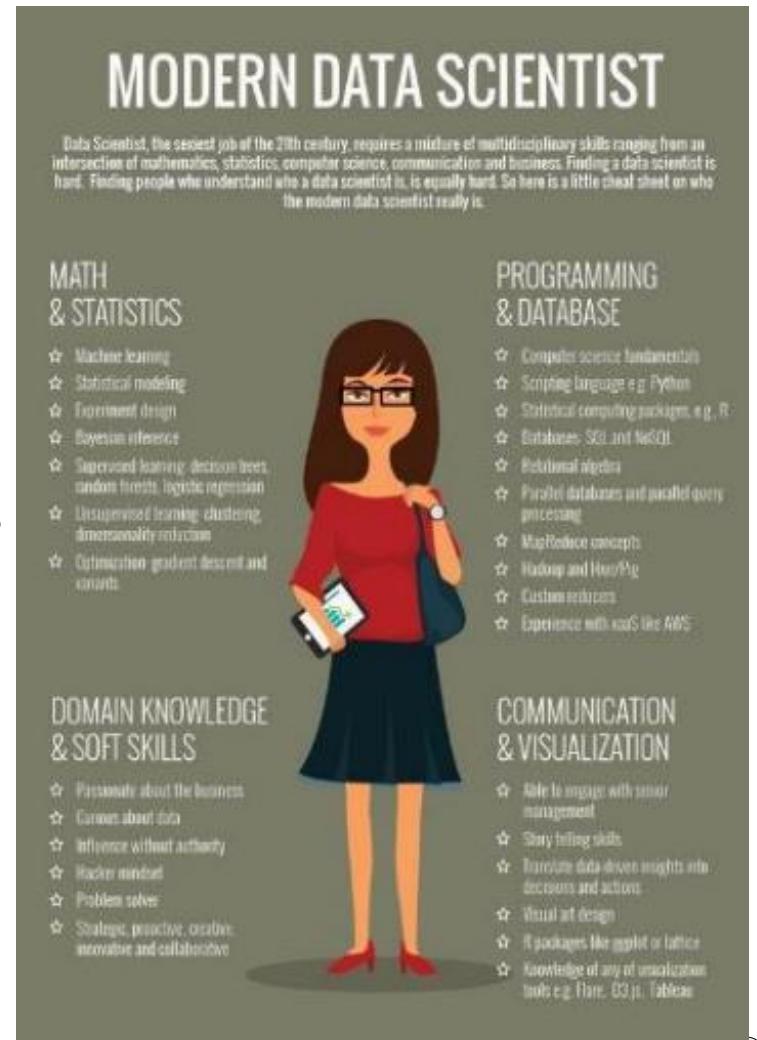
- Analyses / applications / ...
- Documentés en terme de démarche : ce qui a marché / pas marché ...

Des recommandations sur les sources de données utilisées

- Acquisition des données
- Gestion IT
- Backlog des difficultés rencontrées

Des pistes de travail sur de nouvelles exploitations des données

... Des demandeurs qui ont compris le fonctionnement d'un projet data



Exemple de levier Recruter un chief data officer

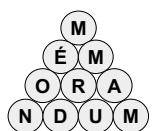
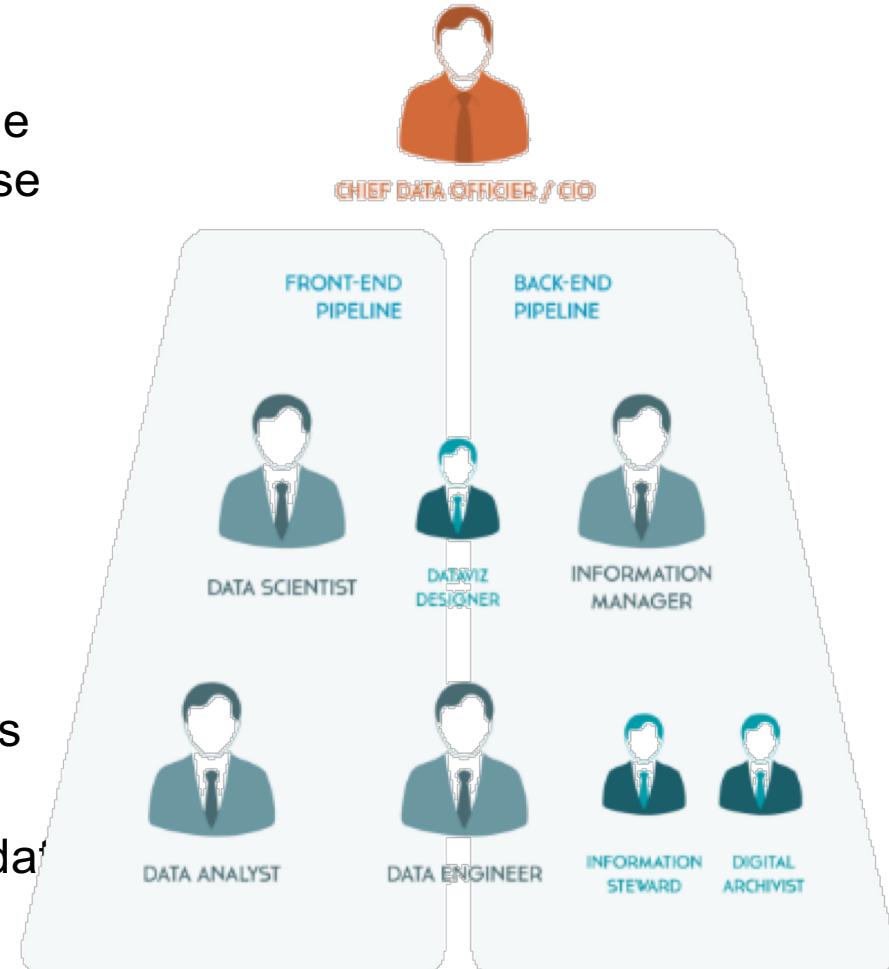
Data science : une ressource nouvelle, justifiant une entité à part au démarrage mais ayant vocation à se dissoudre dans l'entreprise ?

Faut-il les ranger :

- dans l'IT
- Au sein du marketing (et de tout autre direction)
- Dans leur propre tour

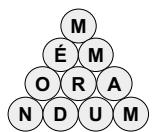
Le CDO répond à d'autres enjeux que manager des data scientists

- Définir une politique de données (interne / open data / ..) sur les volets sécurité
- Faire émerger de nouveaux usages / besoins
- ...



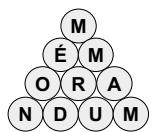
D'autres leviers ?

Séquence crowdsourcing



Animer les autres profils ?

Faire comprendre l'intérêt des données



La difficulté principale est souvent plus dans la question que dans la réponse

Trouver la bonne question à poser à un data scientist est aussi difficile que de choisir les questions qu'on pourrait poser à un Génie

– Espace infini de possibilités, impossibles à hiérarchiser, difficiles à formuler précisément

Heureusement avec un data scientist vous pouvez (et même devez) tester vos questions et les adapter au fur et à mesure que vous obtenez des réponses.

