La Boîte à outils

de la

Stratégie big data

Romain Rissoan
Romain Jouin

DUNOD

Maquette de couverture : Caroline Joubert
Photo de la boîte : © Mega Pixel
Pictos de couverture :
© bioraven-Shutterstock.com
© Eliricon from Noun Project
© I Putu Kharismayadi from Noun Project

Mise en page : Belle Page

© Dunod, 2018
11 rue Paul Bert, 92240 Malakoff
www.dunod.com

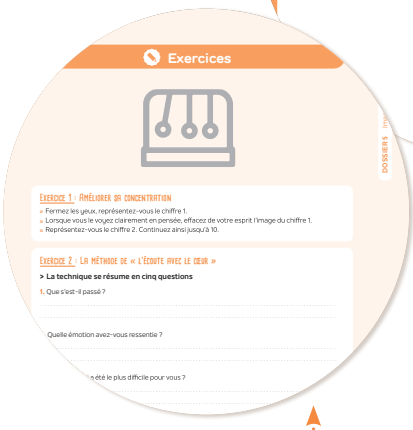
978-2-10-077898-0

VOUS AUSSI, AYEZ LE RÉFLEXE

Boîte à outils

Des outils classés par dossiers thématiques

Une présentation visuelle de chaque outil



Des exemples, cas ou exercices pour approfondir

La Boîte à outils

DES OUTILS OPÉRATIONNELS TOUT DE SUITE

MEGA Boîte à Outils

Manager leader - 100 outils
Coordonnée par Pascale Bélorgey, Nathalie Van Laethem

Digital - 100 outils
Coordonnée par Catherine Lejealle

MÉTIERS

Acheteur, *3e éd.*
Stéphane Canonne, Philippe Petit

Assistante, *2e éd.*
Christine Harache, Hélène Tellitocci

Auditeur financier, *2e éd.*
Sylvan Boccon-Gibod, Eric Vilmint

Chef de produit, *2e éd.*
Nathalie Van Laethem, Stéphanie Moran

Chef de projet, *2e éd.*
Jérôme Maes, François Debois

Coach en entreprise, *2e éd.*
Belkacem Ammiar, Omid Kohneh-Chahri

Commercial, *3e éd.*
Pascale Bélorgey, Stéphane Mercier

Community Manager
Clément Pellerin

Comptabilité, *2e éd.*
Bruno Bachy

Consultant, *2e éd.*
Patrice Stern, Jean-Marc Schoettl

Contrôle de gestion
Caroline Selmer

Création d'entreprise, *2018*
Catherine Léger-Jarniou, Georges Kalousis

E-commerce
Christian Delabre

Formateurs, *3e éd.*
Fabienne Bouchut, Isabelle Cauden, Frédérique Cuisiniez

Management, *2e éd.*
Patrice Stern, Jean-Marc Schoettl

Micro-entrepreneur
Jacques Hellart, Caroline Selmer

Pilote des systèmes d'information, *2e éd.*
Jean-Louis Foucard

Publicité
Servanne Barre, Anne-Marie Gayrard-Carrera

Responsable communication, *3e éd.*
Bernadette Jézéquel, Philippe Gérard

Responsable financier, *2e éd.*
Caroline Selmer

Responsable marketing omnicanal, *3e éd.*
Nathalie Van Laethem, Béatrice Durand-Mégret

Responsable qualité, *3e éd.*
Florence Gillet-Goinard, Bernard Seno

Ressources humaines, *2e éd.*
Annick Haegel

Santé - Sécurité - Environnement, *3e éd.*
Florence Gillet-Goinard, Christel Monar

TPE
Guillaume Ducret

COMPÉTENCES TRANSVERSALES

Conduite du changement
David Autissier, Jean-Michel Moutot

Créativité, *2e éd.*
François Debois, Arnaud Groff, Emmanuel Chenevier

Design management
Bérangère Szostak, François Lenfant

Développement durable et RSE
Vincent Maymo, Geoffroy Murat

Gestion des conflits
Jacques Salzer, Arnaud Stimec

Innovation, *2e éd.*
Géraldine Benoit-Cervantes

Intelligence collective
Béatrice Arnaud, Sylvie Caruso-Cahn

Intelligence économique
Christophe Deschamps, Nicolas Moinet

Lean
Radu Demetrescoux

Leadership, *2e éd.*
Jean-Pierre Testa, Jérôme Lafargue, Virginie Tilhet-Coartet

Management de la relation client, *2e éd.*
Laurence Chabry, Florence Gillet-Goinard, Raphaëlle Jourdan

Management transversal
Jean-Pierre Testa, Bertrand Déroulède

Marketing digital
Stéphane Truphème, Philippe Gastaud

Mind mapping
Xavier Delengaigne, Marie-Rose Delengaigne

Mon parcours professionnel
Florence Gillet-Goinard, Bernard Seno

Négociation, *2e éd.*
Patrice Stern, Jean Mouton

Organisation, *2e éd.*
Benoît Pommeret

Prise de décision
Jean-Marc Santi, Stéphane Mercier, Olivier Arnould

Réseaux sociaux, *4e éd.*
Cyril Bladier

Sécurité économique
Nicolas Moinet

Stratégie, *2e éd.*
Bertrand Giboin

Stratégie digitale omnicanale
Catherine Headley, Catherine Lejealle

Supply chain
Alain Perrot, Philippe Villemus

DÉVELOPPEMENT PERSONNEL

Bien-être au travail
Clothilde Huet, Gaëlle Rohou, Laurence Thomas

Confiance en soi
Annie Leibovitz

Développement personnel en entreprise
Laurent Lagarde

Efficacité professionnelle
Pascale Bélorgey

Gestion du stress
Gaëlle du Penhoat

Gestion du temps
Pascale Bélorgey

Intelligence émotionnelle
Martine-Eva Launet, Céline Peres-Court

Marketing de soi
Nathalie Van Laethem, Stéphanie Moran

Motivation
Sophie Micheau-Thomazeau, Laurence Thomas

Pleine conscience au travail
Sylvie Labouesse, Nathalie Van Laethem

Avant-propos

> Un bon data scientist est intéressé par la résolution de problèmes, pas par de nouveaux outils.
>
> KDNuggets

L'ouvrage que nous vous proposons est dédié au domaine de la data. Il a pour ambition de dresser un panorama des notions, concepts et outils les plus éprouvés sur le marché actuellement. Sans être technique, il vous permettra de connaître les cas d'usage des différents outils et comment ils se distinguent les uns des autres.

Le monde de la data

En 2012, la *Harvard Business Review* (*HBR*) publiait un article : « Data-scientist, the sexiest job of the 21st century ». Depuis, tout le monde veut être data scientist : le gouvernement français a fait de la data science un des piliers de la « nouvelle France industrielle » en 2015, la BPI finance à tour de bras des start-up proposant du machine learning ou du big data et la récente étude « France IA » montre à quel point la France a une carte à jouer dans ce nouveau domaine de l'analyse des données, mais que la partie n'est pas gagnée.
Ce que la *HBR* avait oublié de nous dire, c'est que data scientist est un travail compliqué. Un data scientist doit avoir 3 compétences :

1. informatique
2. mathématique
3. business

Complétées par une excellente capacité à communiquer, c'est-à-dire transmettre ces trois compétences majeures à un auditoire qui ne les a pas.
Ce contexte rend la tâche du data scientist quasi impossible à atteindre. C'est pourquoi nous travaillons en équipe, avec chacun ses compétences. Avec ce livre, vous apprendrez le vocabulaire nécessaire au travail en équipe indispensable dans le monde de la data. Vous pourrez ainsi mettre vos compétences particulières à disposition de vos collègues.

Ce livre est avant tout un livre sur la « stratégie big data » : aussi, ses quatre premiers dossiers s'attachent à expliciter les enjeux qui se cachent derrière ce phénomène :

- Qu'est-ce qui a changé ces dernières années vis-à-vis de la data ?
- Quels sont les différents types de data ?
- Comment en tirer parti pour la stratégie de son entreprise ?
- Quels impacts techniques ?

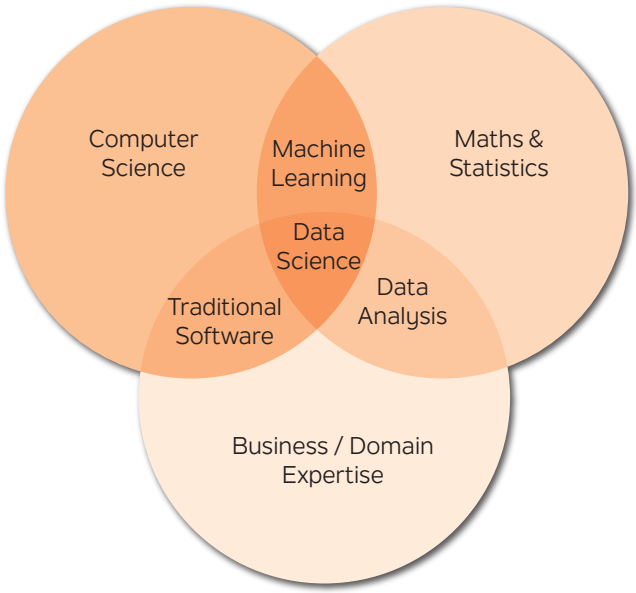
Ces dossiers intéresseront plutôt le manager avide de comprendre comment profiter de cette opportunité qu'est le big data.

Viennent ensuite trois dossiers plus opérationnels, allant de la gestion de projets data à l'introduction des concepts de machine learning, et la différence avec les concepts de big data. Les personnes récemment nommés chef de projets dans un environnement big data y trouveront les clefs pour communiquer et comprendre les équipes avec lesquelles elles devront travailler.

Le dernier dossier, sur le passage en production, est une introduction aux différentes technologies usuellement mises en place dans les infrastructures big data. Il permet d'en comprendre les concepts, sans rentrer dans le détail technique opérationnel en tant que tel.

Ainsi, les différents chapitres sont relativement indépendants, et s'adressent à un public de managers, ou chefs de projet, curieux du phénomène « big data ».

Classiquement, la data science nécessite six expertises distinctes, que l'on cherche chez le data scientist. Évidemment, il est bien difficile de disposer de l'ensemble, c'est pourquoi les projets sont généralement effectués en équipe. Il faut alors s'assurer de partager le même vocabulaire pour que le projet avance correctement.



Les compétences des data scientists

Sommaire