

# Data science: des projets d'un nouveau type ?

Société Suisse de Management de Projet  
09 février 2021

Xavier Bays  
Data Scientist & Cofounder @ Swiss-SDI



# Sommaire

1. Contextualisation
  1. Swiss-SDI
  2. Qu'est-ce que la data science ?
2. Spécificités
  1. Le déroulement d'un projet data
  2. Risques et clés du succès
3. Elaboration et partage
  1. Le business data science canvas
4. Take home message
  1. 3 éléments essentiels

# 1. Contextualisation



# L'expertise de Swiss-SDI

## Conseil

### Conseil stratégique



Audit de la **stratégie IoT**



**Suivi et accompagnement**  
d'initiatives en data science

### Conseil **analytique**



Analyse de la **qualité**  
d'une chaîne de  
production

N!buxs

Ingénieurs civils spécialisés en construction routière

Analyse de la **détérioration**  
des routes suisses

## Développement

### Outils d'interprétation

**LIEBHERR**

Analyse **prédictive** sur les  
moteurs

**FONDAREX**  
SWISS VACUUM TECHNOLOGY

Nouveau dashboard  
intégré à la production

### Algorithmes **de décision**

groupe

**Prédiction** de la consommation  
d'énergie

tpf

Tirage au sort et  
**optimisation** de  
parcours

# Data Science Pourquoi est-ce difficile à définir?



Data Scientist  
Data Engineer  
Data Analyst



Analyse de données  
Data Cleaning  
Features engineering  
Business Intelligence



Intelligence artificielle (I.A.)  
Machine Learning  
Deep Learning  
Reinforcement Learning



Big Data  
Data Lake  
IoT  
Cloud computing

# Data Science: des définitions

*La science des données est l'**extraction de connaissance** d'ensembles de données.*

- Vasant Dhar, Data Science and Prediction

*La science de la donnée, permet **collecter, nettoyer, analyser et tirer parti de données** afin d'en dégager des tendances ou enseignements.*

- Cyrielle Chauwin, Saagie

*La science des données permet à une entreprise d'explorer et d'analyser les données brutes pour les transformer en **information précieuses** permettant **désoudre les problèmes de l'entreprise***

- Bastien L., Le Big Data

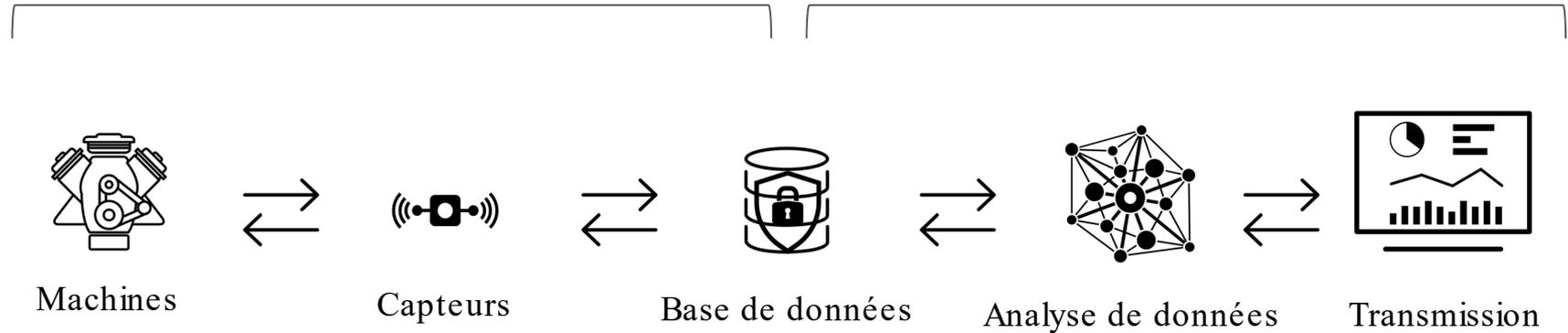
Selon Jim Gray, la science des données est le **quatrième paradigme de la science**

- empirique,
- théorique,
- computationnelle,
- data-driven.

# Data Science: les éléments clés

Flux / Collecte

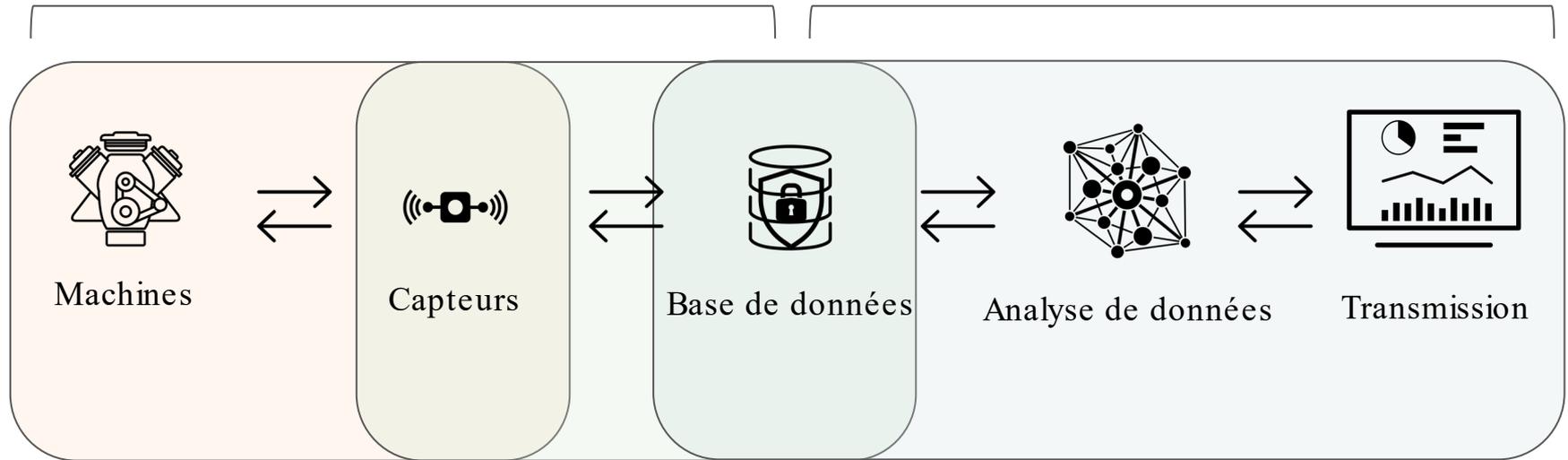
Analyse / Actionnables



# Data Scientist ou Data Engineer ?

Flux / Collecte

Analyse / Actionnables

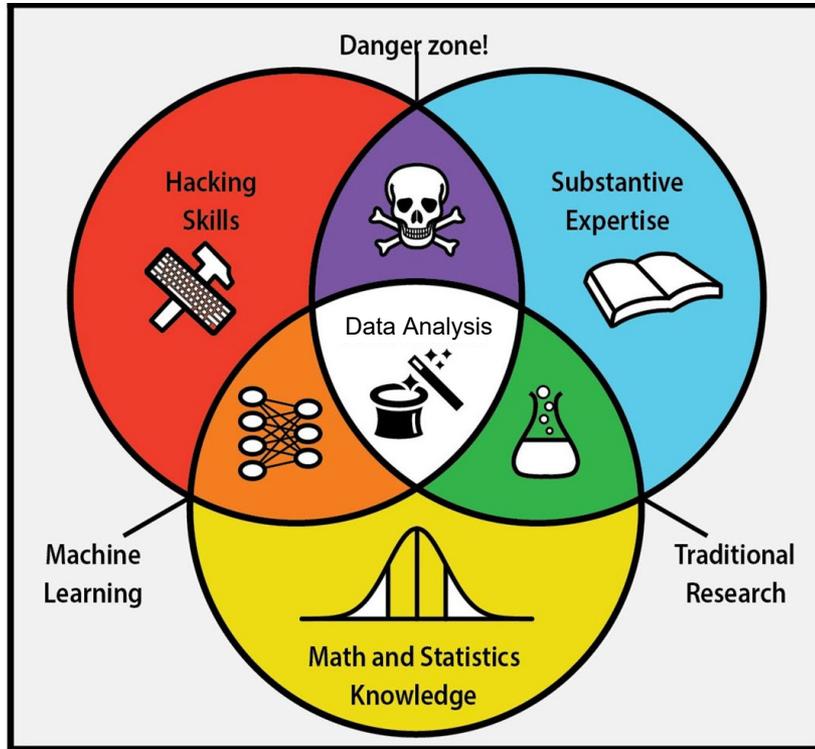


Télémétrie

Data Engineer

Data Scientist

# L'analyse de données



## Mathématiques & Statistiques

- Connaissances des modèles sous-jacents
- Capacités de créer des modèles sur mesure

## Programming Skills

- Connaissances des outils de programmation
- Capacités dans la gestion des données

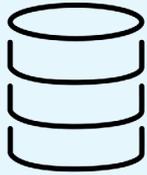
## Expertise

- Connaissances des besoins de l'entreprise ou du domaine
- Capacités de transmission

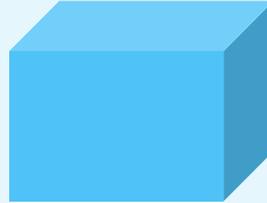
# Machine Learning le cerveau de l'I.A.

## Phase d'Apprentissage

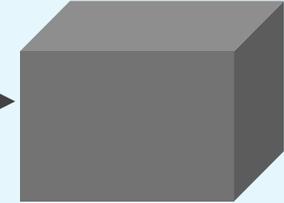
Données  
D'apprentissage



Algorithme  
Ignorant

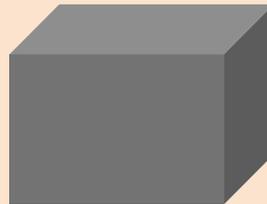
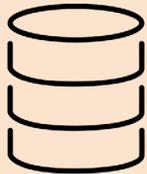


Algorithme  
Intelligent



## Phase de Prédiction

Nouvelles  
Données



Prediction



# Le Big Data: un contexte

## Définition basique :

Quand un ordinateur standard ne suffit plus !

## Définition des V's:

### Définitions des 3 V's:

- Volume
- Variété
- Vitesse

Business

### Définitions des 5 V's:

- Volume
- Variété
- Vitesse
- **Véracité**
- **Valeur**

# La data science: Quelques exemples... des médias!



Recommandations personnalisées



## Netflix

### This is how Netflix's top-secret recommendation system works

Netflix splits viewers up into more than two thousands taste groups. Which one you're in dictates the recommendations you get

## BuzzFeed News

REPORTING TO YOU

ABOUT US GOT A TIP? SUPPORT US BUZZFI



TECH

## Netflix Wants To Change The Way You Chill

Log on to Netflix today, and you'll see a lot of stuff you've never heard of. Here's how the streaming service is convincing you to click on it anyway.



**Nicole Nguyen**  
BuzzFeed News Reporter

Posted on December 13, 2018, at 3:45 p.m. ET



# La data science: Quelques exemples... des médias!



Recommandations personnalisées



Traducteurs automatiques

The screenshot shows the RTVE website interface. At the top, there is a navigation bar with categories like 'Noticias', 'Televisión', 'Radio', 'Deportes', 'Infantil', 'A la carta', 'El Tiempo', and 'Playz'. Below this, a secondary navigation bar lists various news topics such as 'Datos Coronavirus', 'La gran vacunación', 'Elecciones Cataluña', 'VerificaRTVE', 'Telediario en 4', 'Radio 5 +24', 'España', 'Mundo', 'Economía', 'Cultura', 'Tecnología', and 'Más temas'. The main content area features a news article titled 'Google Translate, seis años traduciendo palabras y 200 millones de usuarios'. The article includes a sub-header '» Noticias » Ciencia y tecnología' and a list of bullet points: 'La compañía puso en marcha su traductor original en 2006', 'Actualmente realiza traducciones entre 64 lenguas diferentes', and 'El 92% de usuarios del servicio procede de fuera de Estados Unidos'. Below the article, there is a section for 'RTVE.es / EUROPA PRESS' and a 'LesEchos' logo. At the bottom, there is a navigation bar with links for 'À la une', 'Idées', 'Économie', 'Politique', 'Monde', 'Tech-Médias', 'Entreprises', 'Bourse', 'Finance - Marchés', 'Régions', 'Patrimoine', and 'Le Mag W-E'. There are also icons for 'En direct', 'Le Journal', 'Newsletters', and 'Podcasts', along with a 'CONNEXION' button.

## Comment l'IA révolutionne la traduction

L'histoire de l'intelligence artificielle est intimement liée à celle de la traduction, peut-être parce que savoir donner du sens est ce qu'il y a de plus humain et de plus mystérieux pour la machine.



# La data science. Quelques exemples... des médias!



Recommandations personnalisées



Traducteurs automatiques



Véhicules autonomes



# La data science: Quelques exemples...du terrain!



## Marketing

- Publicité ciblée
- Amélioration de la connaissance des consommateurs



## Finances

- Détection de fraudes
- Trading



## Production

- Maintenance prédictive
- Optimisation des processus
- Contrôle qualité

## 2. Les spécificités



# La nature du projet data

## De la transversalité,

- Des compétences multiples
- Des rencontres inédites
- Une information étonnamment répartie

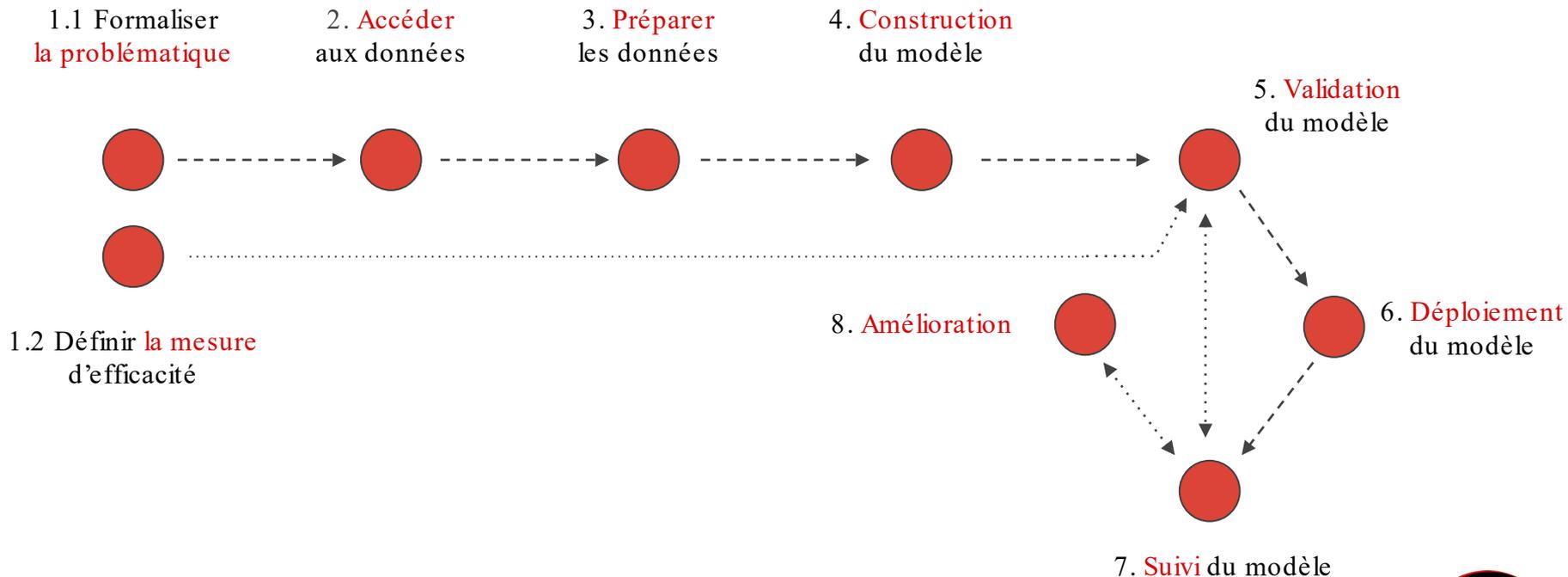
## de l'innovation...

- Des technologies nouvelles
- Le besoin de tester (POC, adoption)
- Des données jamais explorées
- Des questions d'un nouveau type

... et des surprises garanties !

# Workflow d'un projet de data science

Point de vue: **data scientist**



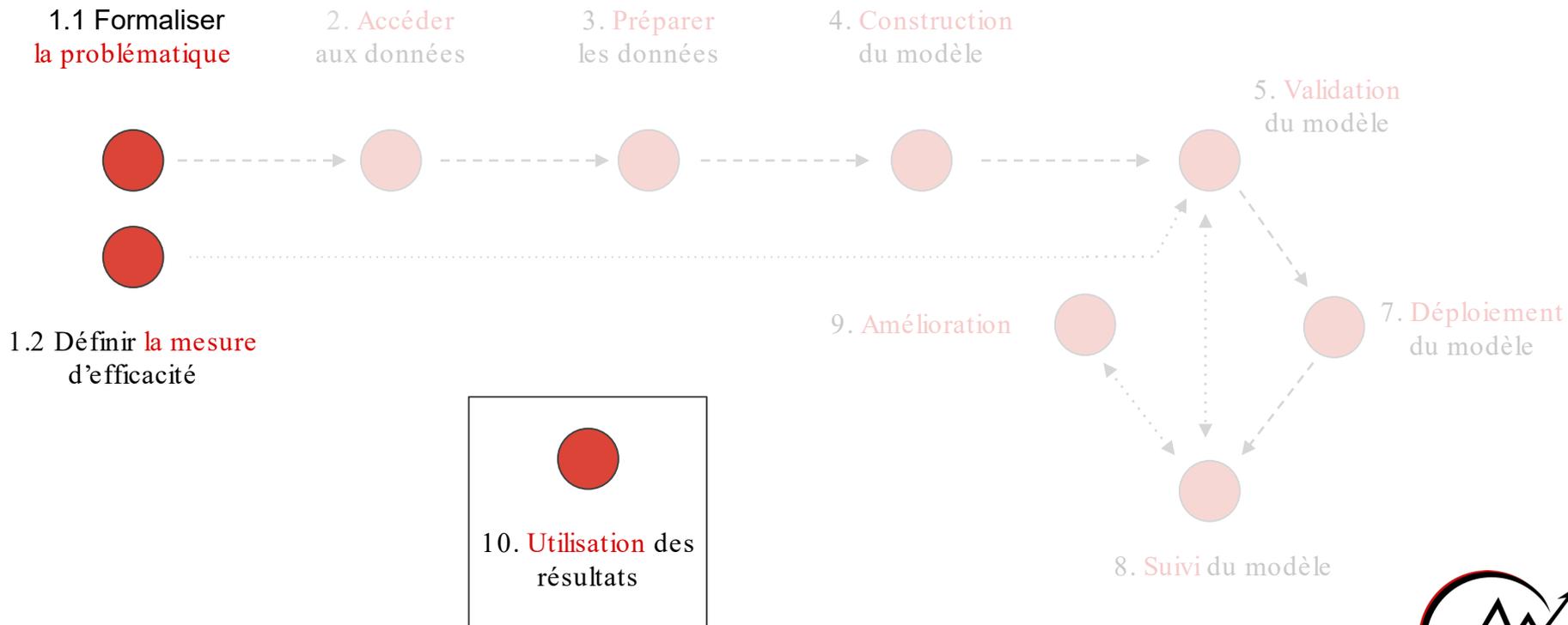
# Workflow d'un projet de data science

Point de vue: gestionnaire de projet

- Elaboration
- POC
- Développement
- Mise en production
- Mise en place de la structure de suivi

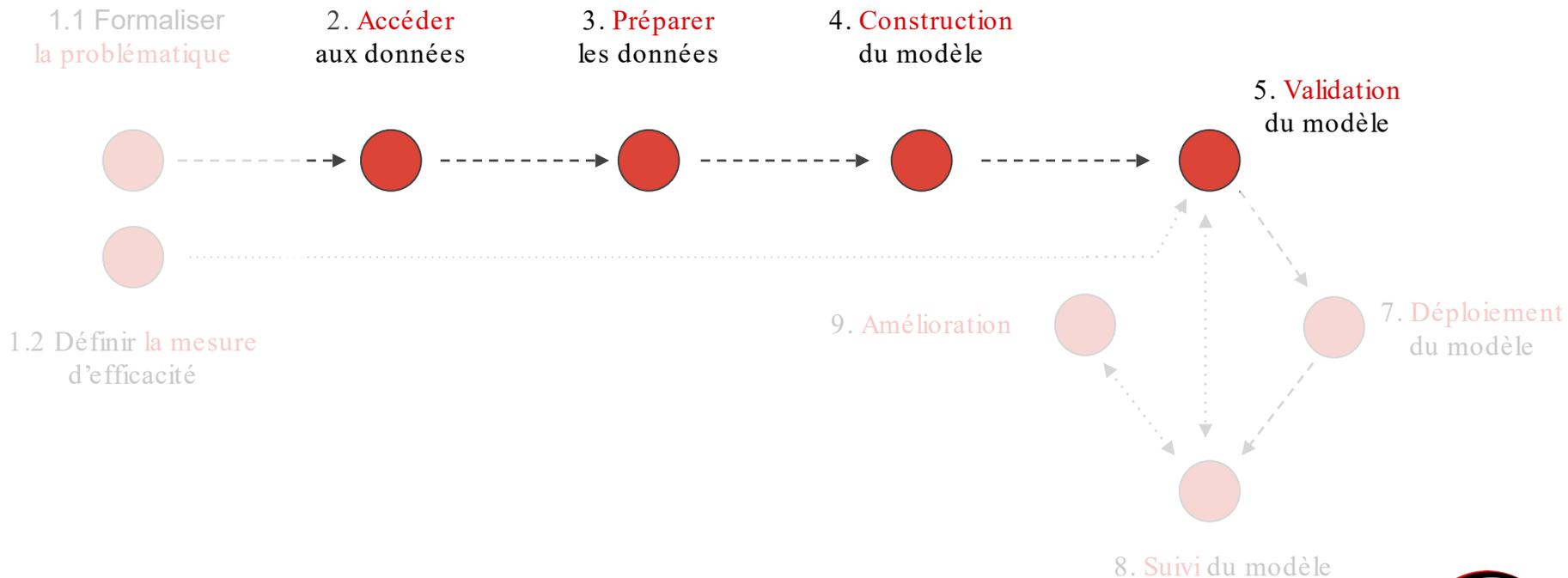
# Workflow d'un projet de data science

## Elaboration



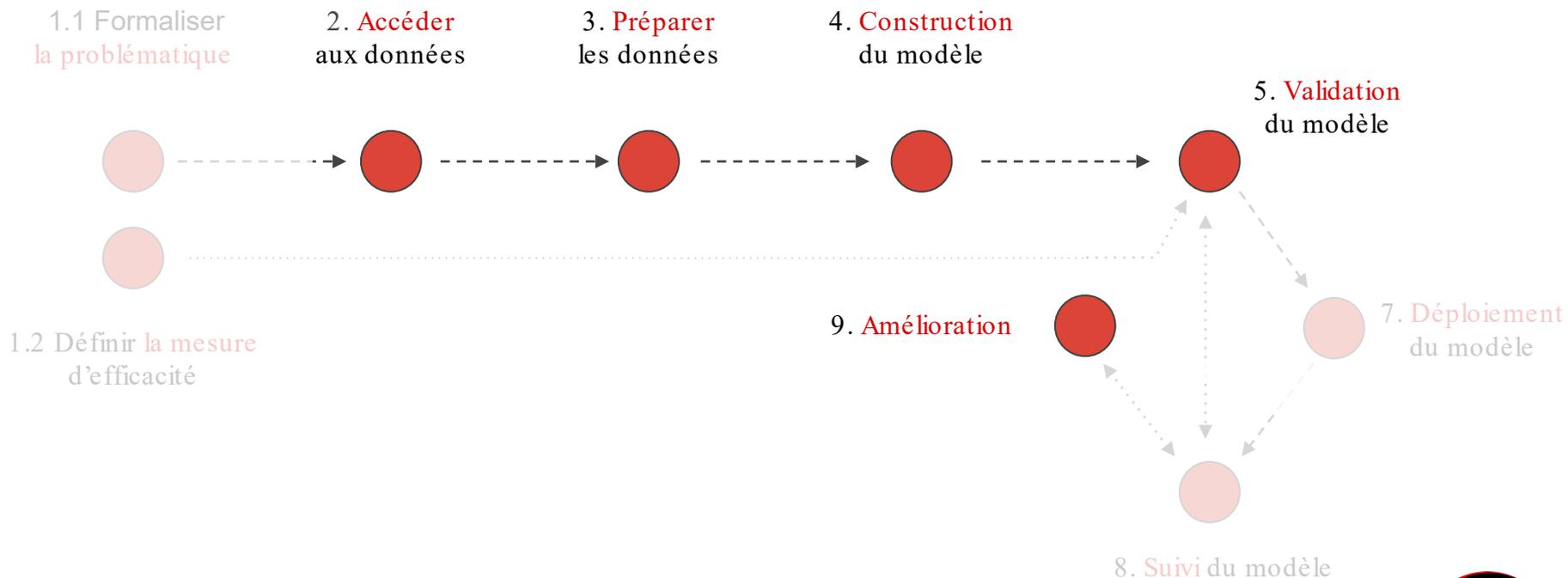
# Workflow d'un projet de data science

## POC



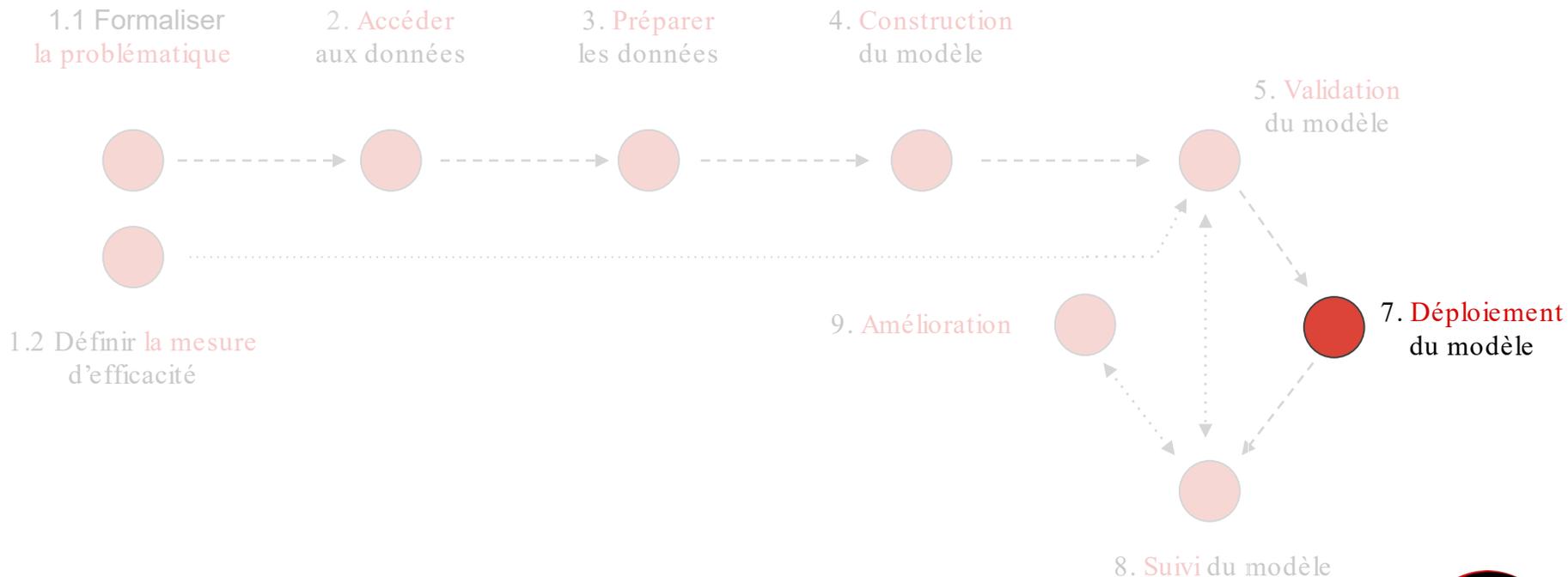
# Workflow d'un projet de data science

## Développement



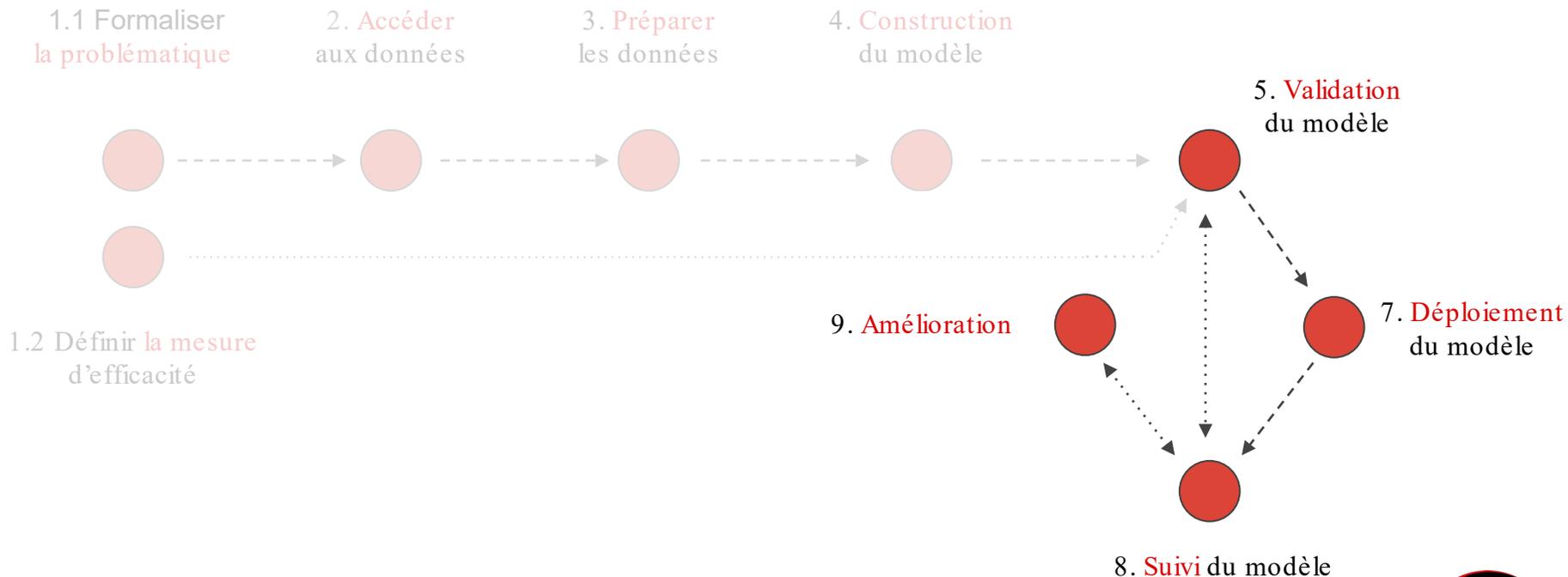
# Workflow d'un projet de data science

## Mise en production



# Workflow d'un projet de data science

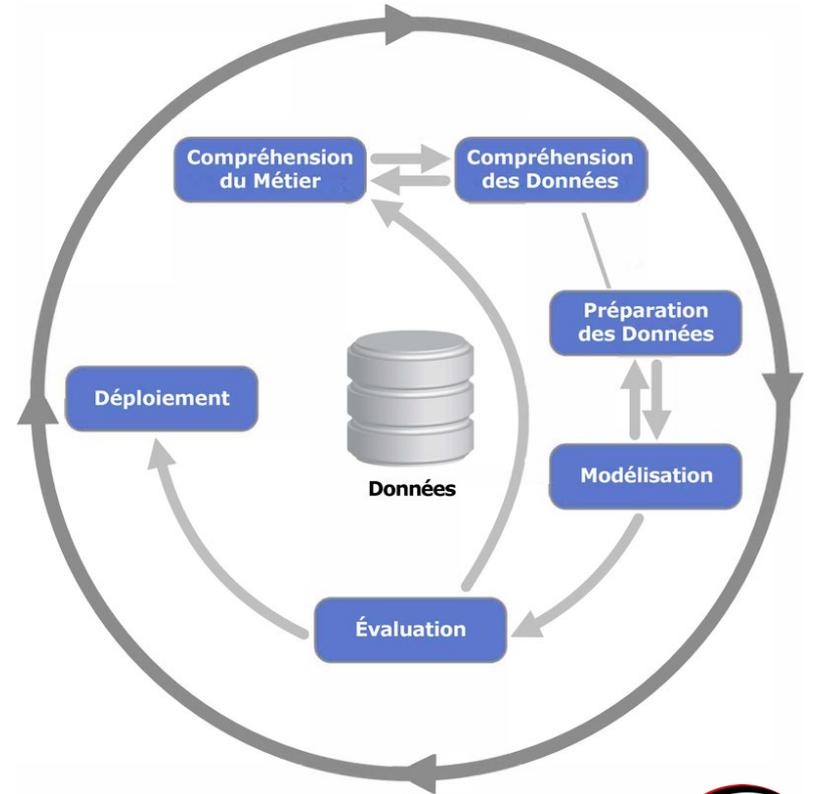
## Mise en place de la structure de suivi



# Les méthodologies existantes

Peu de méthodologies propres

- CRISP-DM
  - cf image ci-contre
- SCRUM
  - Agilité
- Waterfall
  - Phases linéaires et séquentielles
- Kanban
  - Done, In progress, To do
- WBS
  - fractionnement tâche individuelle



# Les Risques

Les risques/ interrogations les plus fréquemment avancés sur le terrain:

- Techniques
  - Qualité des données
  - Biais
  - Mise en production
- Humains
  - Ethique
  - RGPD
  - Adoption
- Conception
  - Exigences irréalistes
  - ROI inexistant

# Les clés du succès

Des éléments à ne pas négliger:

- Objectifs et critère de satisfaction bien définis
- Démarche itérative / agile
- Compréhension des données
- Un soutien global

# 3. Elaboration et partage



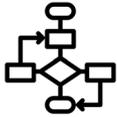
# Pourquoi élaboration et partage?

## Notre quotidien

- Entreprises d'horizons multiples
- Rencontre avec
  - des directeurs
  - des managers en R&D
  - des experts métiers
- Notre rôle
  - Conseiller
  - Designer
  - Mener des projets

# Pourquoi élaboration et partage?

## Les besoins du terrain



Répondre à des problématiques bien définies



Lier data science et business

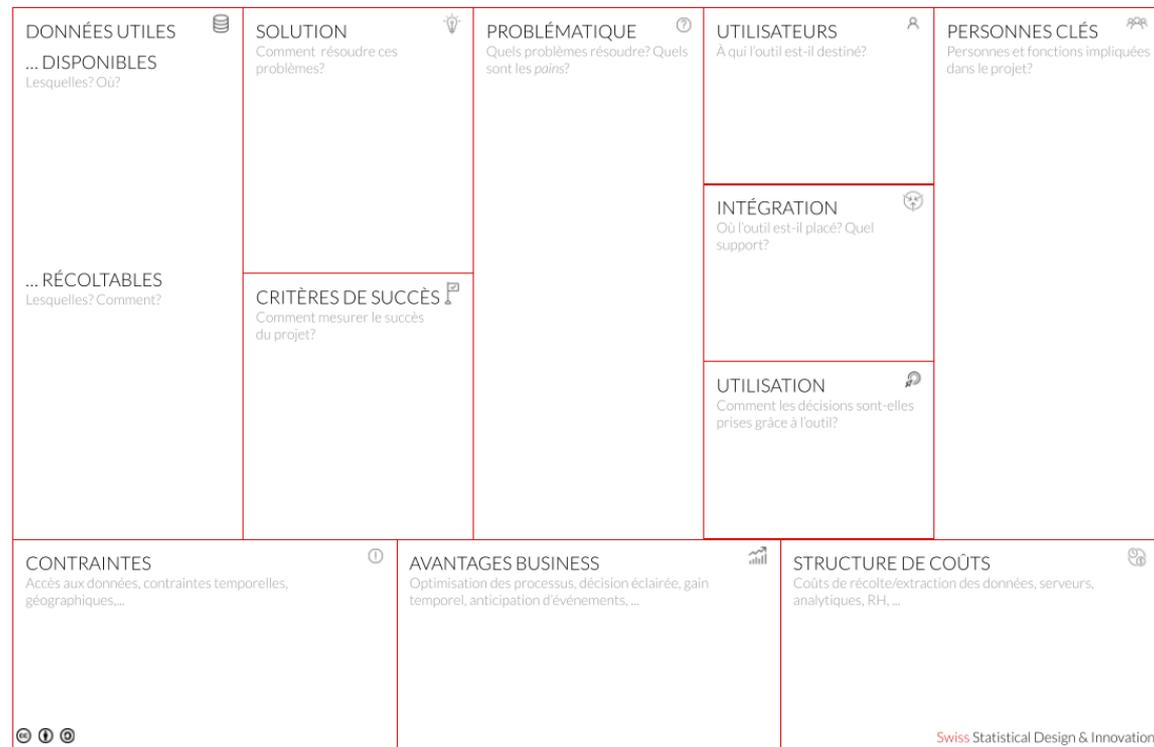


Intégrer les concepts d'interactions humain-machine

# The Business Data Science Canvas (BDS-Canvas)

The Business Data Science Canvas (v1.1)

Titre:



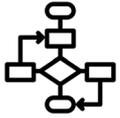
Licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

This work is inspired by the Business model canvas of Osterwalder, Pigneur & al

- **Source d'inspiration**  
Business model canvas  
Osterwalder, Pigneur & al.
- **Objectifs**  
Elaborer des projets de data science en prenant en compte le maximum d'enjeux pratiques.
- **Unique Value Proposition**  
Lie business et data science dans un unique canevas.

# Focus sur leBDS-Canvas

Quatre angles de vue



**Problématique:** partie centrale



**Data:** partie de gauche



**Business:** partie basse

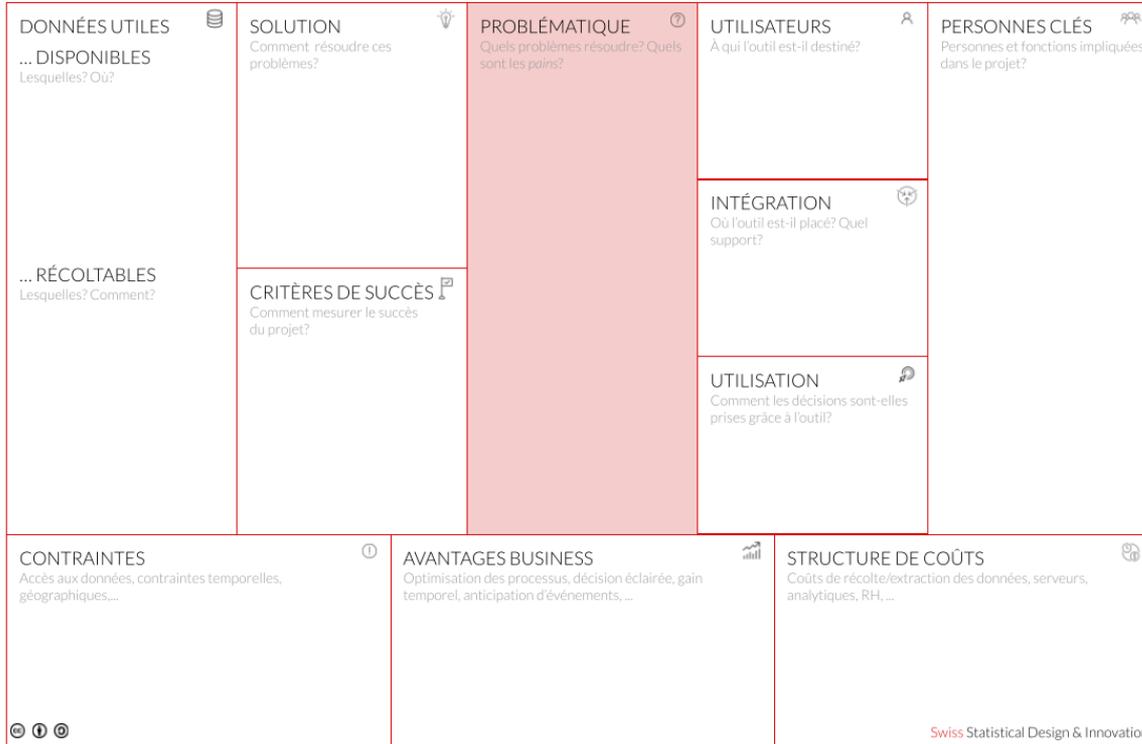


**Humain:** partie de droite

# BDS-Canvas: la problématique

The Business Data Science Canvas (v1.1)

Titre:



- **Problématique bien définie**  
“On veut faire de la maintenance prédictive!”
- **Élément central du Canevas**  
Pas de problème, pas de solution
- **Une modélisation...**  
pour une problématique.

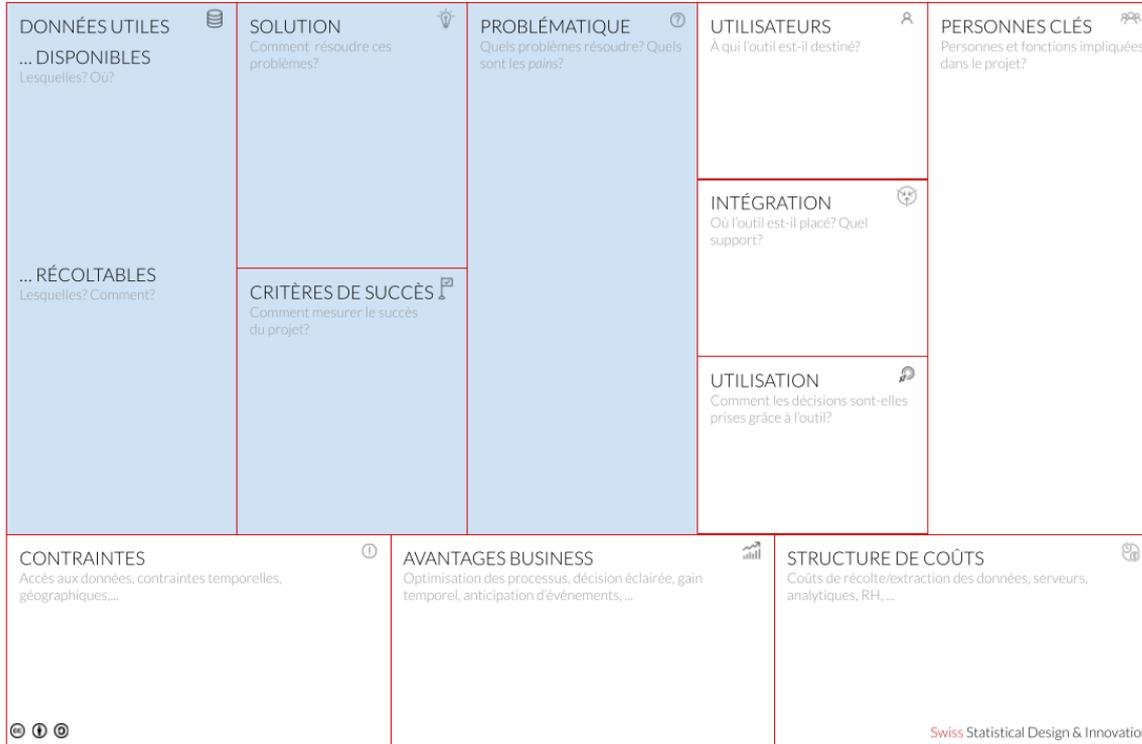
Licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

This work is inspired by the Business model canvas of Osterwalder, Pigneur & al

# BDS-Canvas: focus data

The Business Data Science Canvas (v1.1)

Titre:



- **La solution mathématique** dépend de la formulation de la problématique!
- **Les critères de succès** permettent d'évaluer la qualité du modèle.
- **Les données utiles** sont la matière première du projet.

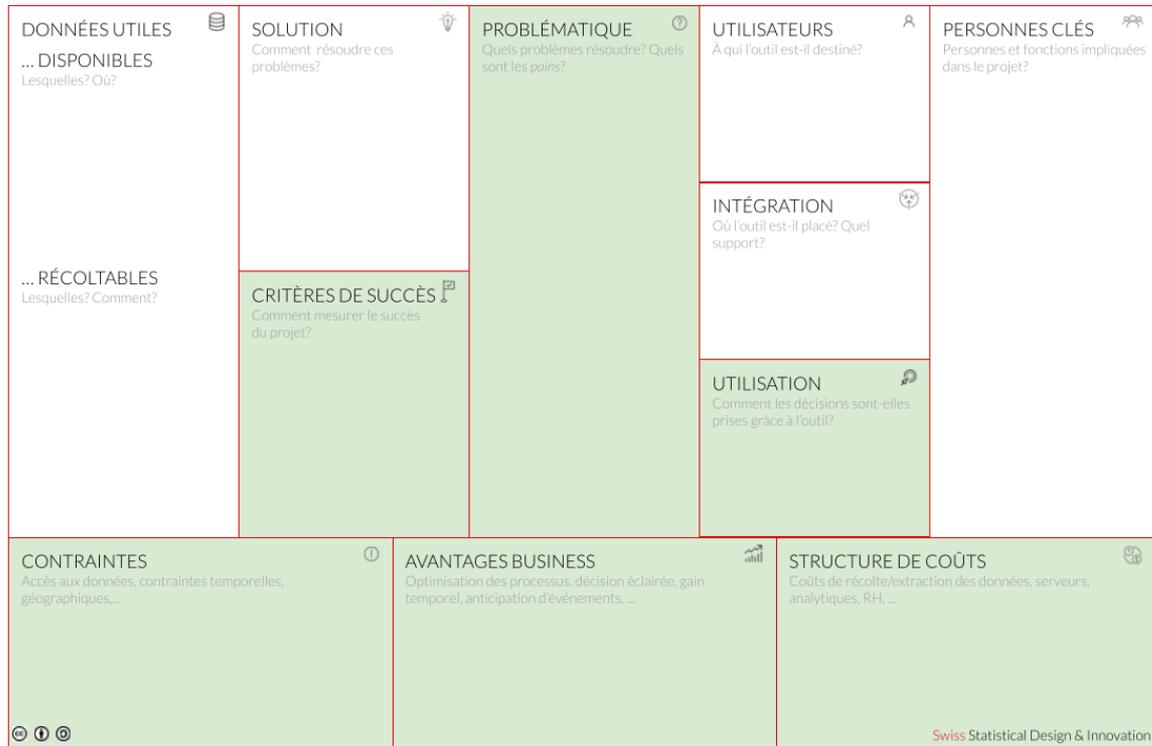
Licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

This work is inspired by the Business model canvas of Osterwalder, Pigneur & al

# BDS-Canvas: focus business

The Business Data Science Canvas (v1.1)

Titre:



Licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

This work is inspired by the Business model canvas of Osterwalder, Pigneur & al

- **Le critère de succès** qualifie la satisfaction de la réalisation du projet
- **L'utilisation de l'outil** permet de prendre des actions concrètes.
- **Les avantages et les coûts** définissent le retour sur investissement.



# 4. Take home message



# L'essentiel en trois points

- Réolvez des problèmes
- Soyez réactifs
- Entourez-vous des bonnes personnes

Merci de votre attention!

Xavier Bays

Data Scientist & Co-fondateur @swissdi

[x.bays@swissdi.ch](mailto:x.bays@swissdi.ch) | + 41 79 530 33 19

Swiss Statistical Design & Innovation Sàrl

Grand-Rue 96A

CH - 1627 Vaulruz

[swiss-sdi.ch](http://swiss-sdi.ch)

