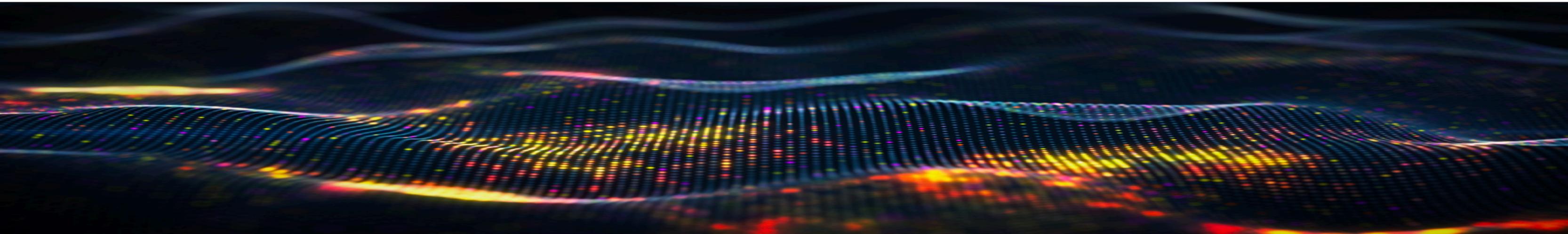


Mise en place de solution de traitement avancé des données

Florian Gaté - Data Scientist





# **Mise en place** de solution de traitement avancé des données

Florian Gaté - Data Scientist



# Mener à bien un projet data

Le besoin

# Mener à bien un projet data



# Mener à bien un projet data



- Les étapes
- Les erreurs à éviter
- Cas d'étude : l'industrie 4.0

# L'industrie 4.0 : Qu'est-ce que c'est ?



1ère Revolution 1765

**Vapeur / mécanisation  
de la production**



# L'industrie 4.0 : Qu'est-ce que c'est ?



1ère Revolution 1765

**Vapeur / mécanisation  
de la production**



2ème Revolution 1870

**Electricité  
Ligne d'assemblage**



# L'industrie 4.0 : Qu'est-ce que c'est ?



1ère Revolution 1765  
**Vapeur / mécanisation  
de la production**



2ème Revolution 1870  
**Electricité  
Ligne d'assemblage**



3ème Revolution 1969  
**Informatique  
Automatisation**



# L'industrie 4.0 : Qu'est-ce que c'est ?



1ère Revolution 1765  
**Vapeur / mécanisation  
de la production**



2ème Revolution 1870  
**Electricité  
Ligne d'assemblage**



3ème Revolution 1969  
**Informatique  
Automatisation**



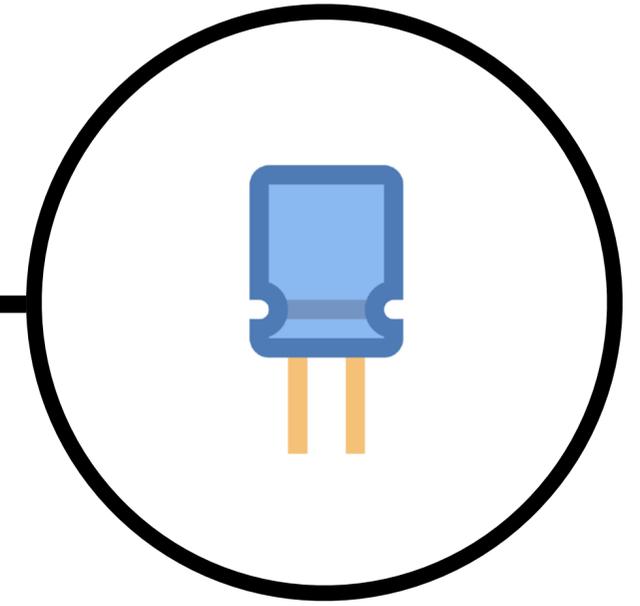
4ème Revolution  
**Traitement de l'information**



## Le cas d'étude



Groupe industriel

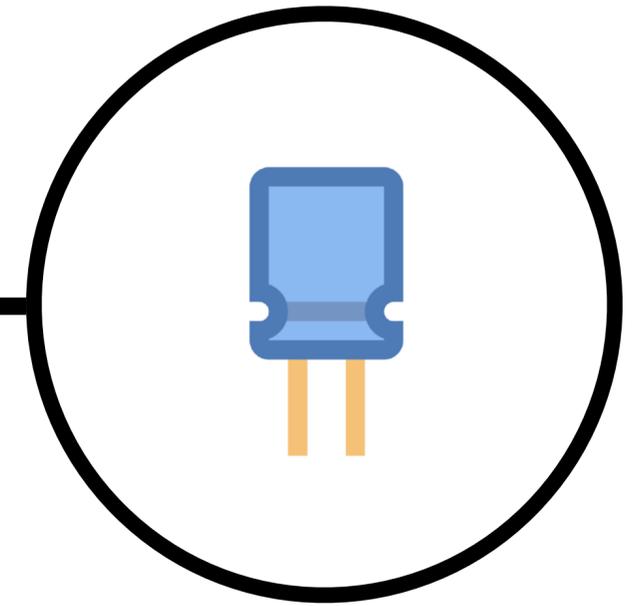


Semi-conducteurs

## Le cas d'étude



Groupe industriel



Semi-conducteurs

- Production automatisée sur lignes d'assemblage
- Certaines machines critiques tombent en panne
- Une panne entraîne l'arrêt de la production

## Prenons un exemple



Groupe industriel

Responsable  
de production

Semi-conducteurs

Problème à résoudre : diminuer les périodes d'arrêt de production

## Prenons un exemple



Groupe industriel

Responsable  
de production

Semi-conducteurs

Problème à résoudre : diminuer les périodes d'arrêt de production

Actuellement : maintenance préventive planifiée

## Prenons un exemple



Groupe industriel

Responsable  
de production

Semi-conducteurs

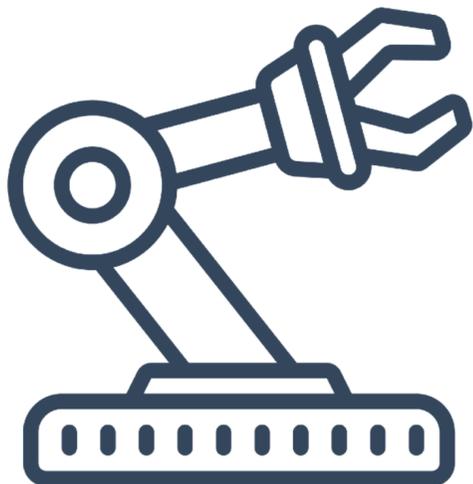
Problème à résoudre : diminuer les périodes d'arrêt de production

**Solution : la maintenance préventive conditionnelle ?**

## Le cas d'étude



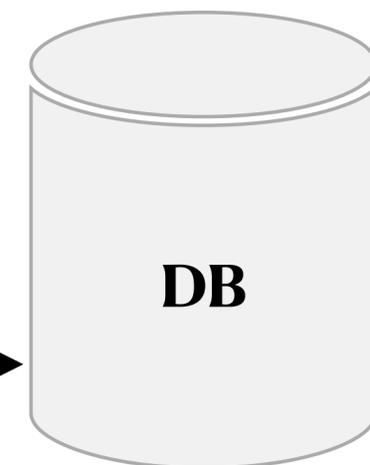
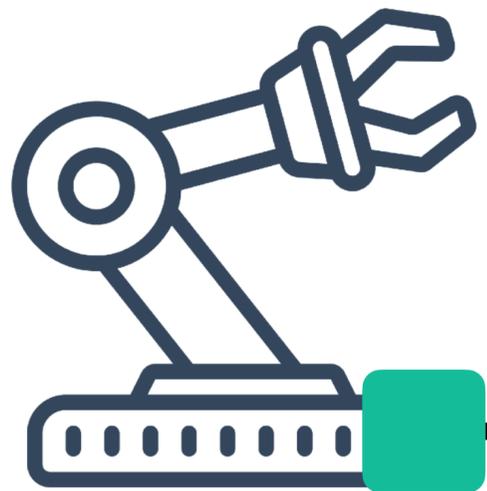
**1ère idée: la maintenance conditionnelle**



## Le cas d'étude



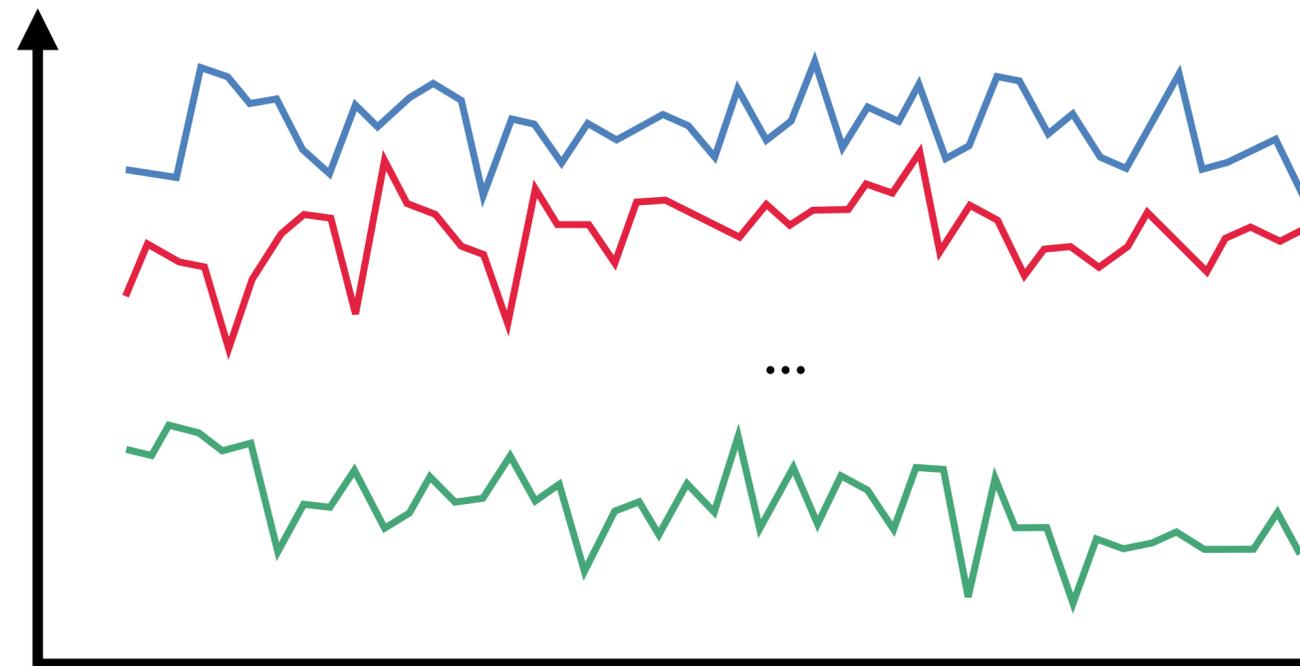
**1ère idée: la maintenance conditionnelle**



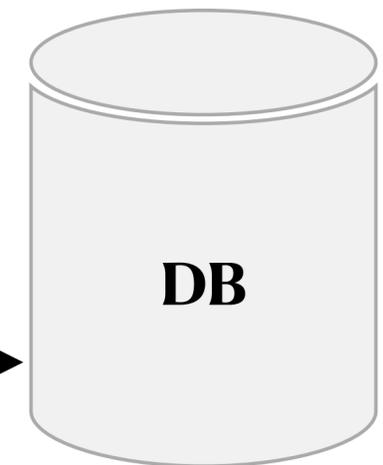
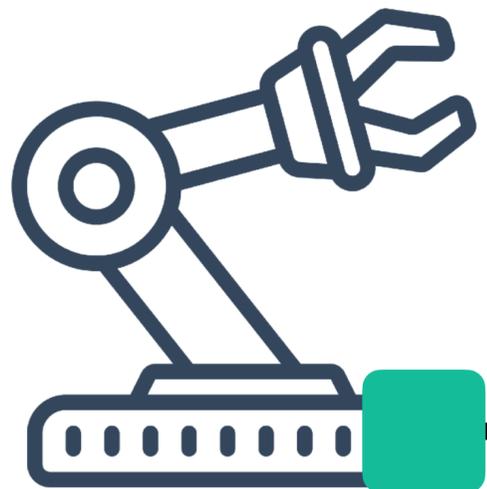
# Le cas d'étude



## 1ère idée: la maintenance conditionnelle



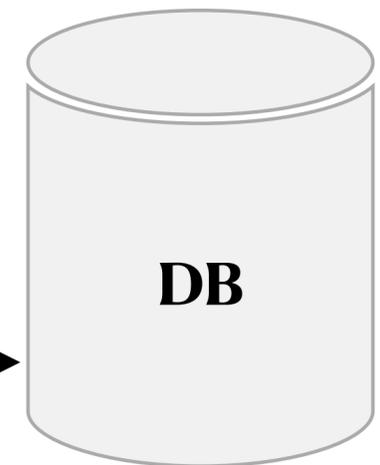
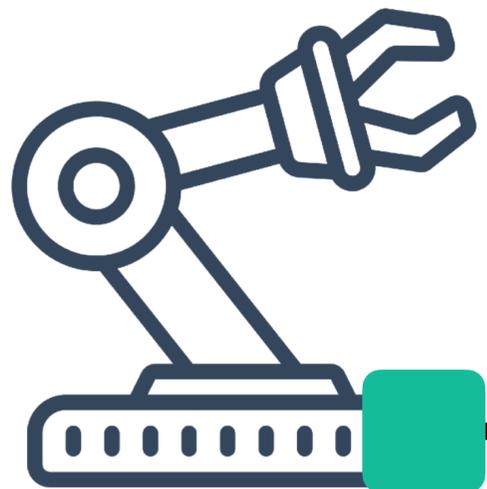
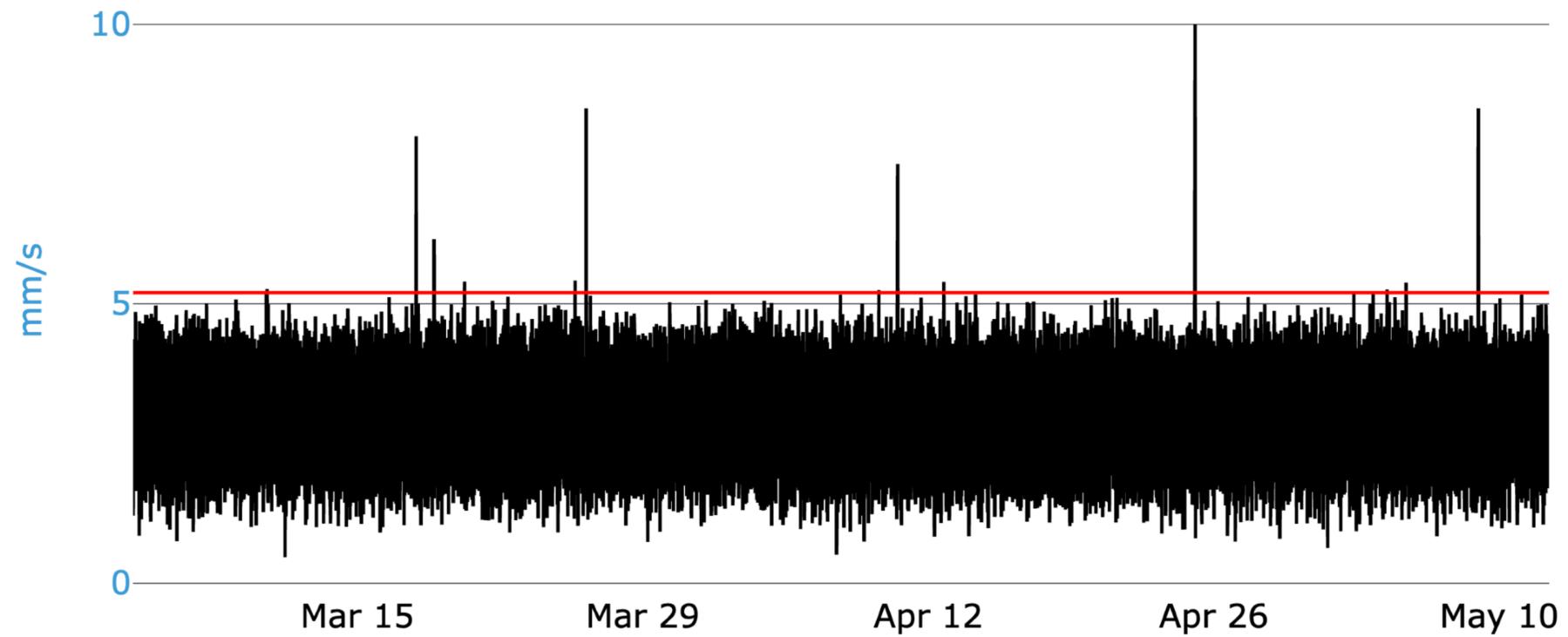
Pression  
Température  
Humidité  
Vibration  
...



# Le cas d'étude



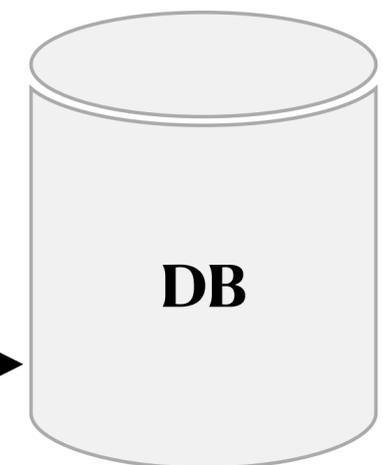
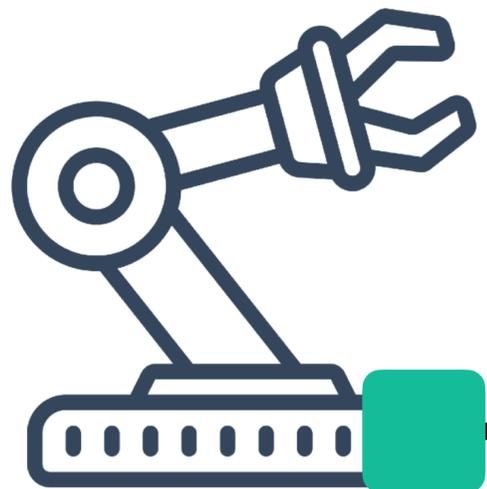
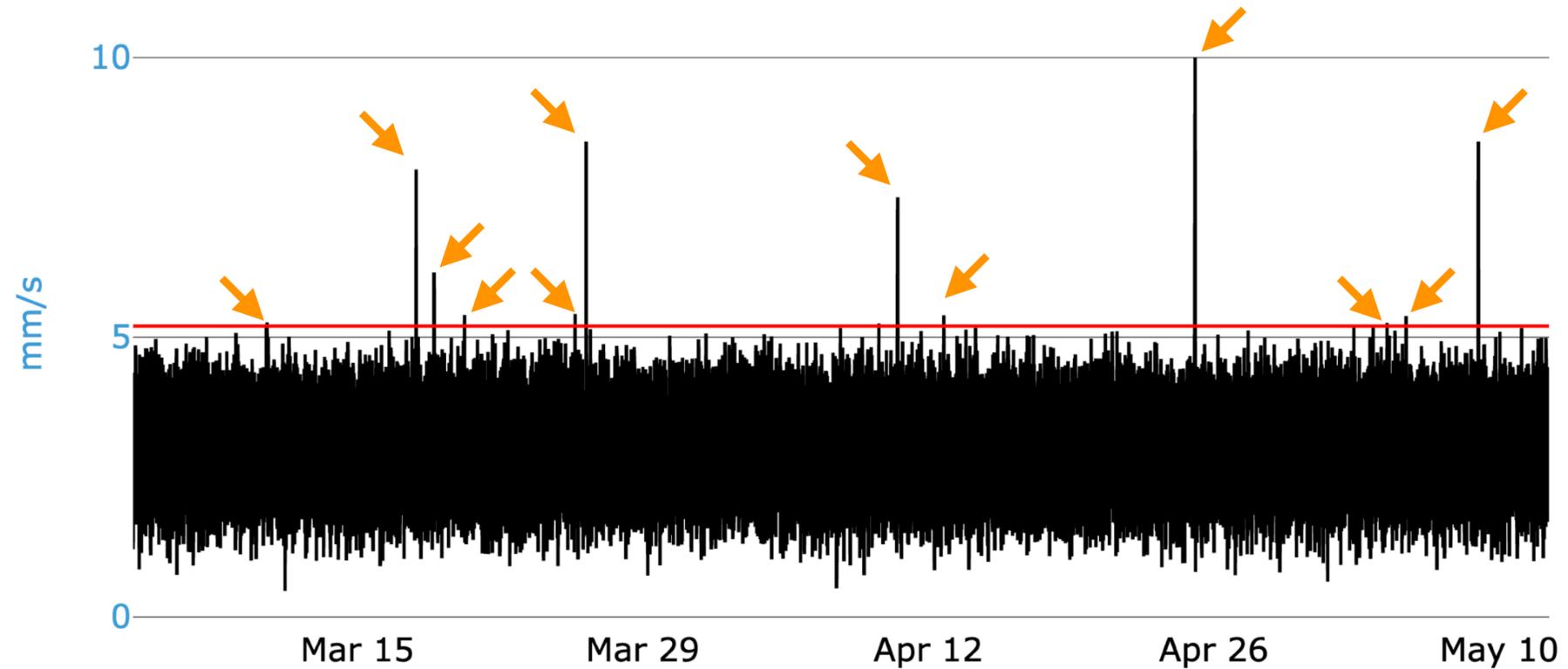
## 1ère idée: la maintenance conditionnelle



# Le cas d'étude



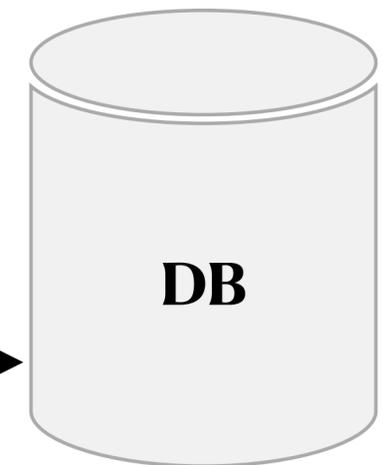
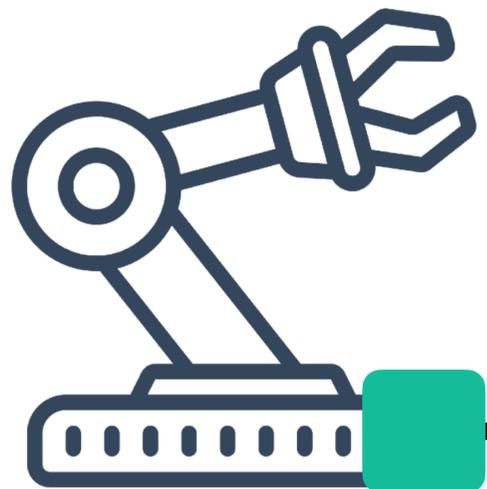
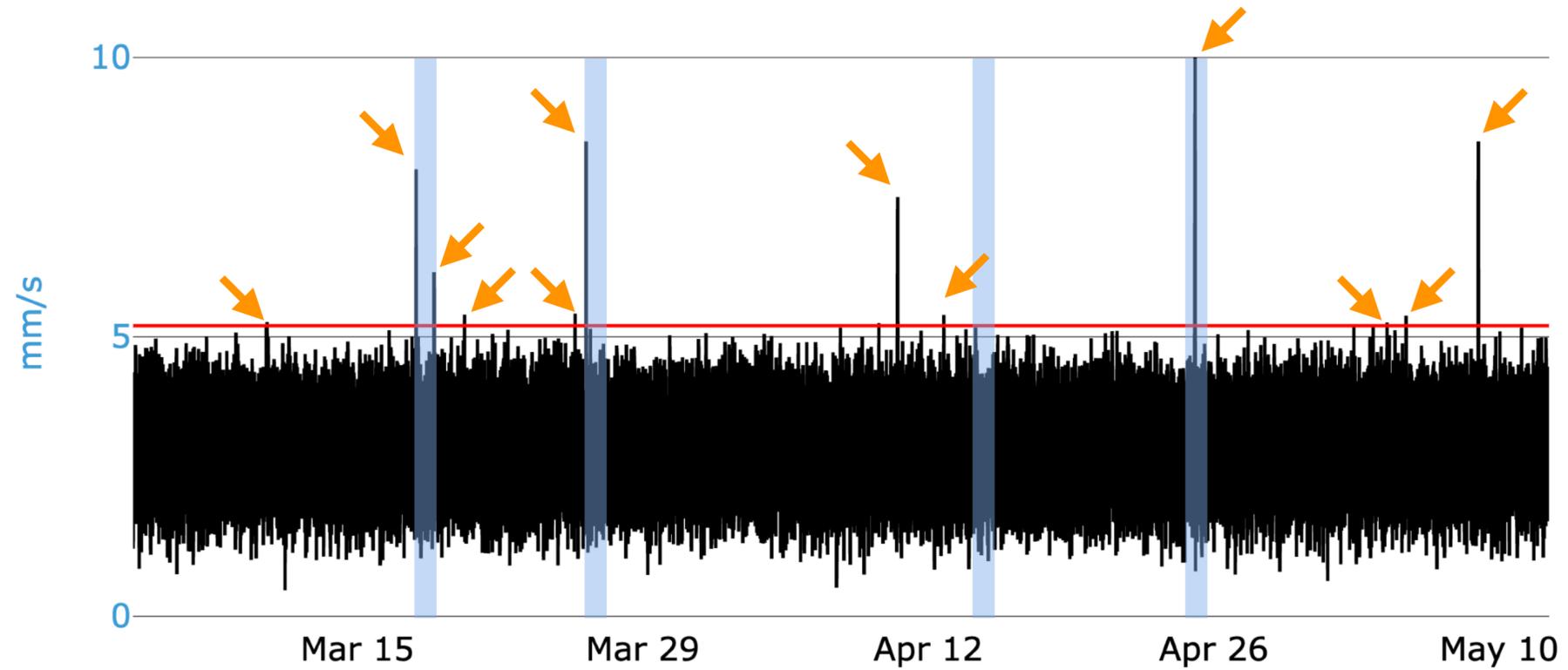
## 1ère idée: la maintenance conditionnelle



# Le cas d'étude



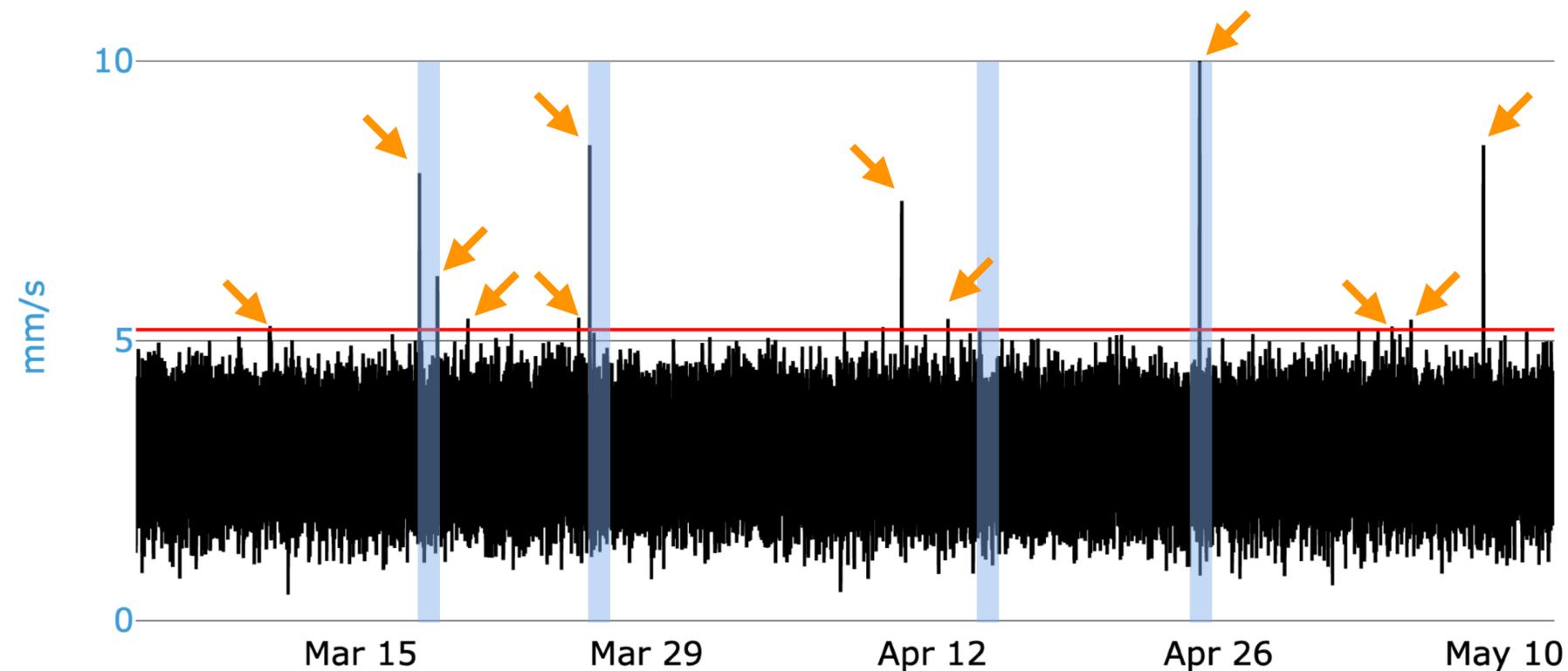
## 1ère idée: la maintenance conditionnelle



# Le cas d'étude

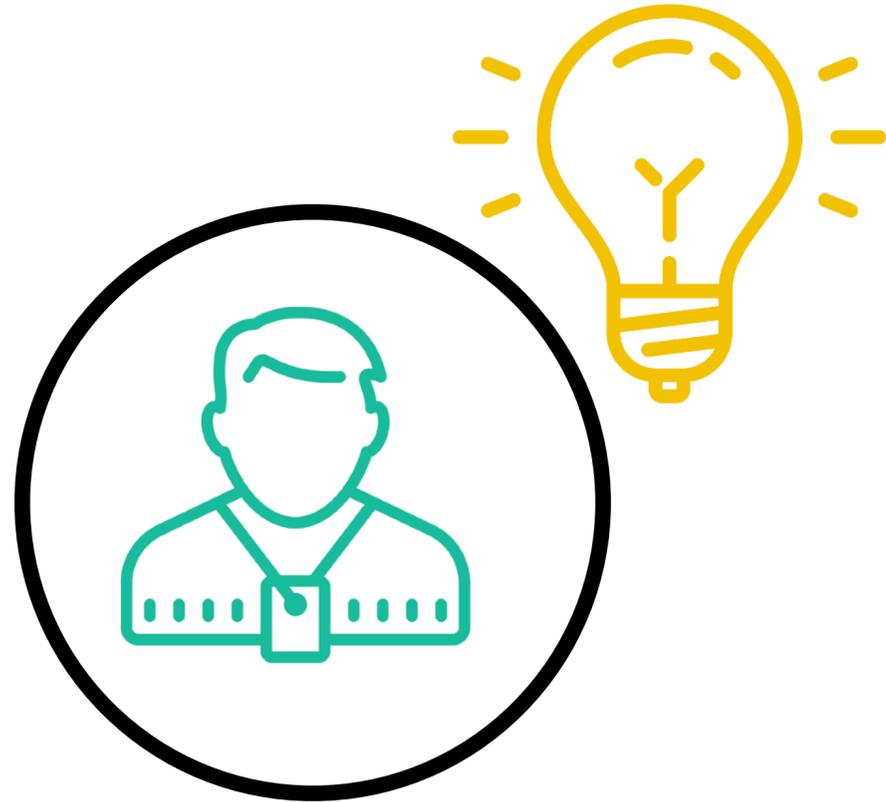


## 1ère idée: la maintenance conditionnelle



- La méthode des seuils sur les grandeurs prises individuellement n'est pas assez précise
- Incidents mal prédits, pas de logique

## Etape 1 : formalisation du besoin

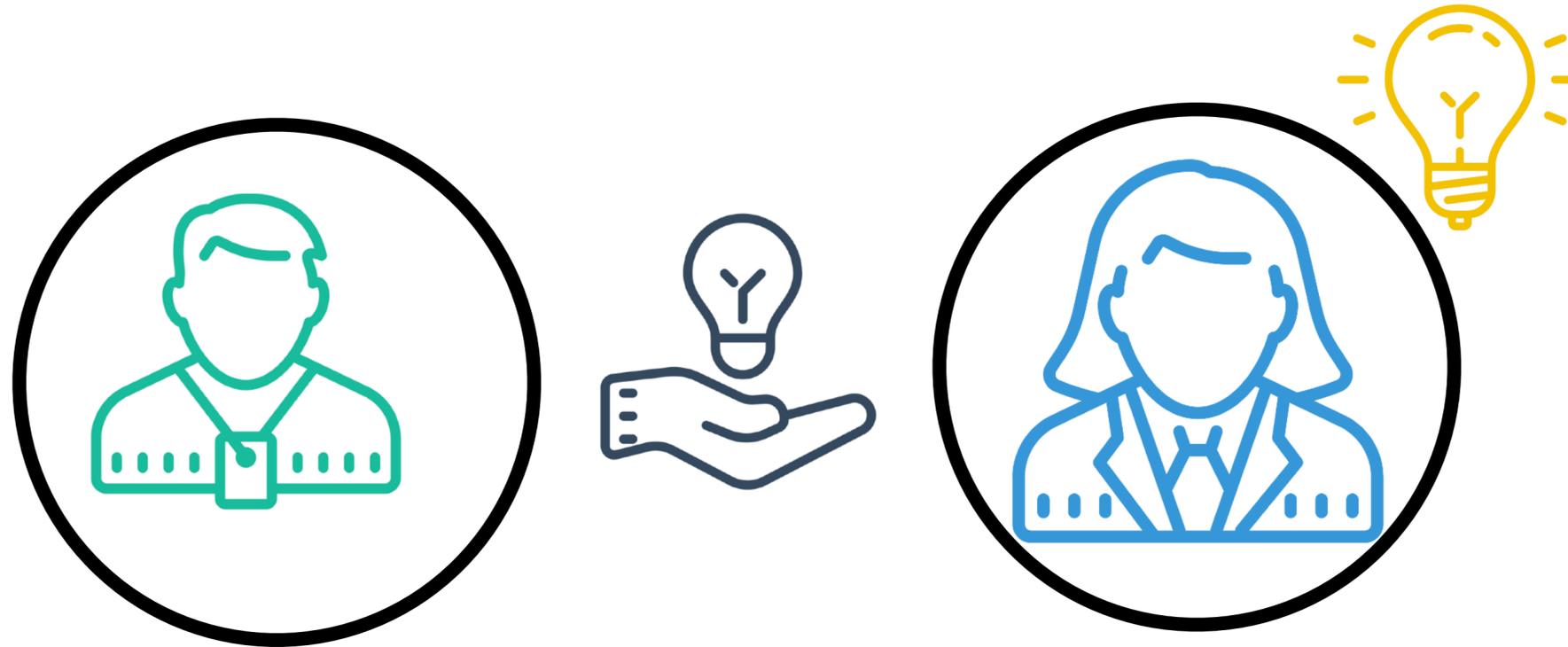


Passer de la maintenance préventive  
à la maintenance **prédictive**

**Besoin** : pouvoir anticiper efficacement les futures défaillances

**Comment** : en analysant les données de manière plus intelligente

## Etape 2 : Convaincre



Avoir le soutien actif de l'équipe dirigeante

Assurer des ressources suffisantes au projet

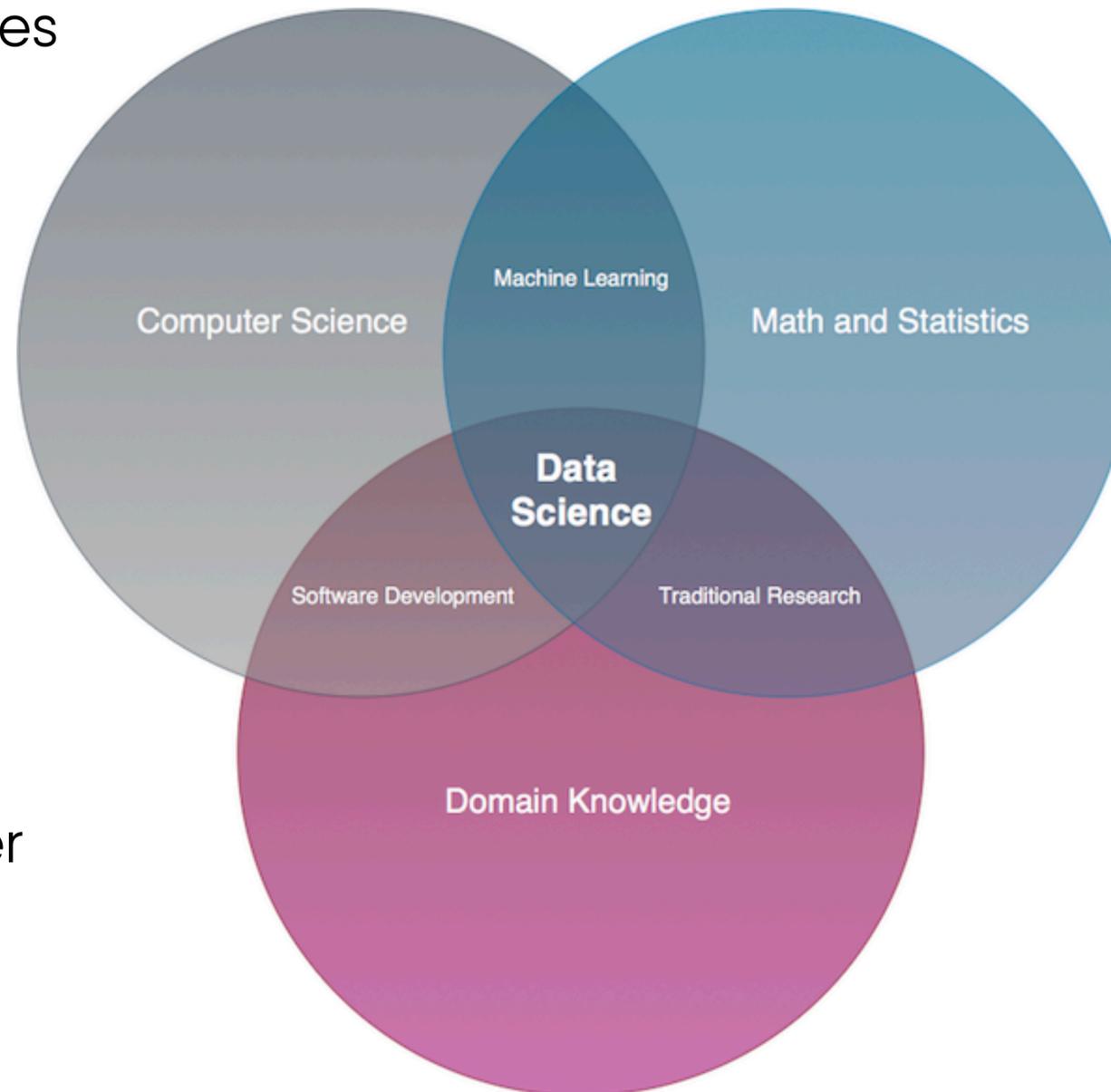
## Etape 3 : Avoir la bonne équipe

### Ingénieur informatique

- Transfert des données
- Stockage
- Disponibilité
- Intégration

### Business

- Connaissance métier
- Contraintes
- Livrables



### Data scientist

- Données nécessaires
- Analyses
- Machine learning

## Etape 4 : Lancement du projet

- Communication entre les différents acteurs
- Traduire les attentes coté business
- Et les solutions proposées coté technique

## Etape 4 : Lancement du projet

- Communication entre les différents acteurs
- Traduire les attentes coté business
- Et les solutions proposées coté technique

## Business

- But : optimiser la charge de travail via les opérations de maintenance
- Moyen : exploiter les données issues des capteurs
- Livrable : web app pour transmettre des alertes automatiques aux techniciens
- ROI : optimisation de production, stockage, couts de maintenance, sécurité, expérience client, ...

## Etape 4 : Lancement du projet

- Communication entre les différents acteurs
- Traduire les attentes coté business
- Et les solutions proposées coté technique

### Business

- But : optimiser la charge de travail via les opérations de maintenance
- Moyen : exploiter les données issues des capteurs
- Livrable : web app pour transmettre des alertes automatiques aux techniciens
- ROI : optimisation de production, stockage, couts de maintenance, sécurité, expérience client, ...

### Ingénieur informatique

- Sources des données
- Accessibilité
- Format
- Architecture (datalake)
- Intégration

## Etape 4 : Lancement du projet

- Communication entre les différents acteurs
- Traduire les attentes coté business
- Et les solutions proposées coté technique

## Data scientist

- Données à dispositions / volume
- Fréquence des pannes
- Pourcentage d'incidents dans les données **historiques**
- Échantillonnage temporel
- Méta données des machines (âges, localisations, ...)
- **Evaluation de solution en fonction de l'historique**

## Etape 4 : Lancement du projet

- Communication entre les différents acteurs
- Traduire les attentes coté business
- Et les solutions proposées coté technique

### Data scientist

- Données à dispositions / volume
- Fréquence des pannes
- Pourcentage d'incidents dans les données **historiques**
- Échantillonnage temporel
- Méta données des machines (âges, localisations, ...)
- **Evaluation de solution en fonction de l'historique**

### Coté Maintenance

- Expérience terrain
- Suggestions causes défaillances
- Autres données ?
- Suggestions livrable

## Etape 4 : Lancement du projet

- Communication entre les différents acteurs
- Traduire les attentes coté business
- Et les solutions proposées coté technique

### **Estimation des coûts de déploiement de la solution**

**Objectif : précision de prédiction des pannes de 20% minimum**

## Etape 5 : La preuve de concept

Elle permet :

- De redéfinir la stratégie
- De redéfinir les objectifs

## Etape 5 : La preuve de concept

Elle permet :

- De redéfinir la stratégie
- De redéfinir les objectifs

Gros avantage :

- Permet d'échouer rapidement
- Donc de limiter les investissements

# Etape 5 : La preuve de concept    Data scientist

## Qualité des données

- Récupération des données disponibles en collaboration avec l'équipe informatique
- Enrichissement
- Nettoyage, traitement des valeurs manquantes, des doublons, ...
- Formattage

# Etape 5 : La preuve de concept    Data scientist

## Qualité des données

- Récupération des données disponibles en collaboration avec l'équipe informatique
- Enrichissement
- Nettoyage, traitement des valeurs manquantes, des doublons, ...
- Formattage

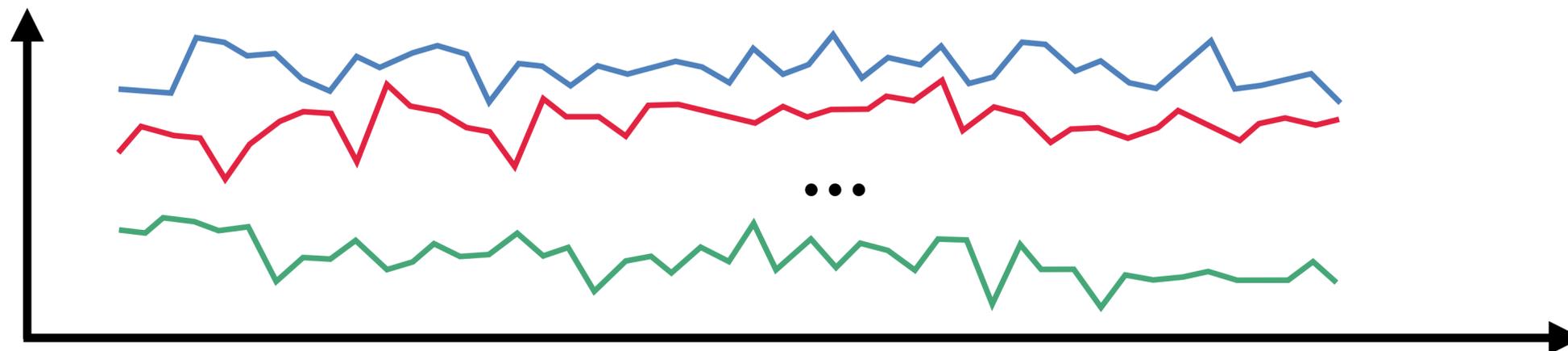
## Phase d'exploration

- Visualisation
- Études de corrélations
- Paramètres d'intérêt
- Comportement symptomatique de défaillances futures ?
- Mise en place du protocole de test

## Etape 5 : La preuve de concept    Data scientist

### Phase de R&D : présentation de la méthode

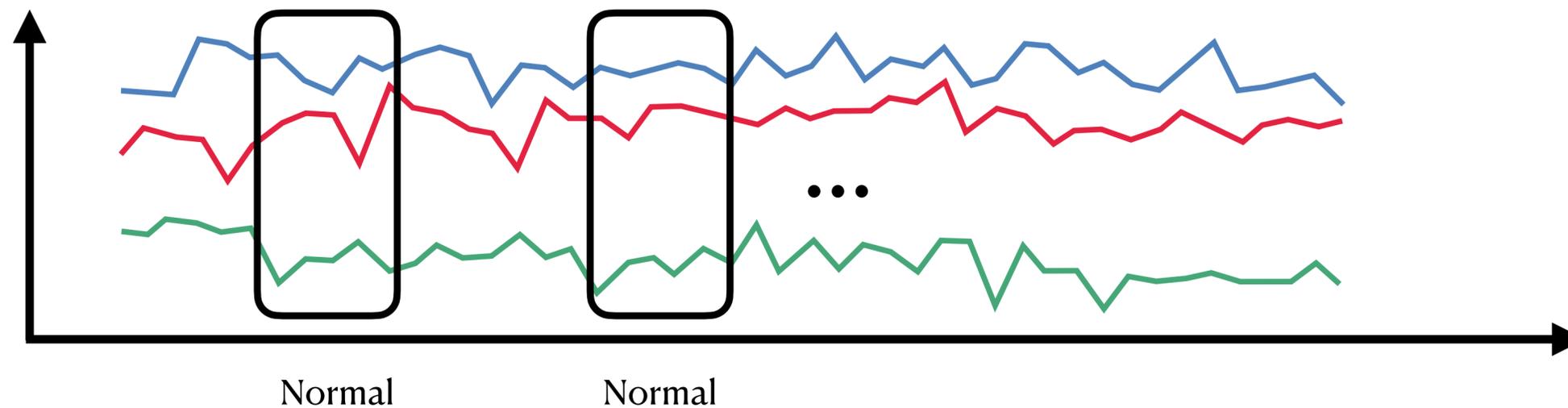
- A priori, pas de différence entre comportement normal et anormal



## Etape 5 : La preuve de concept    Data scientist

### Phase de R&D : présentation de la méthode

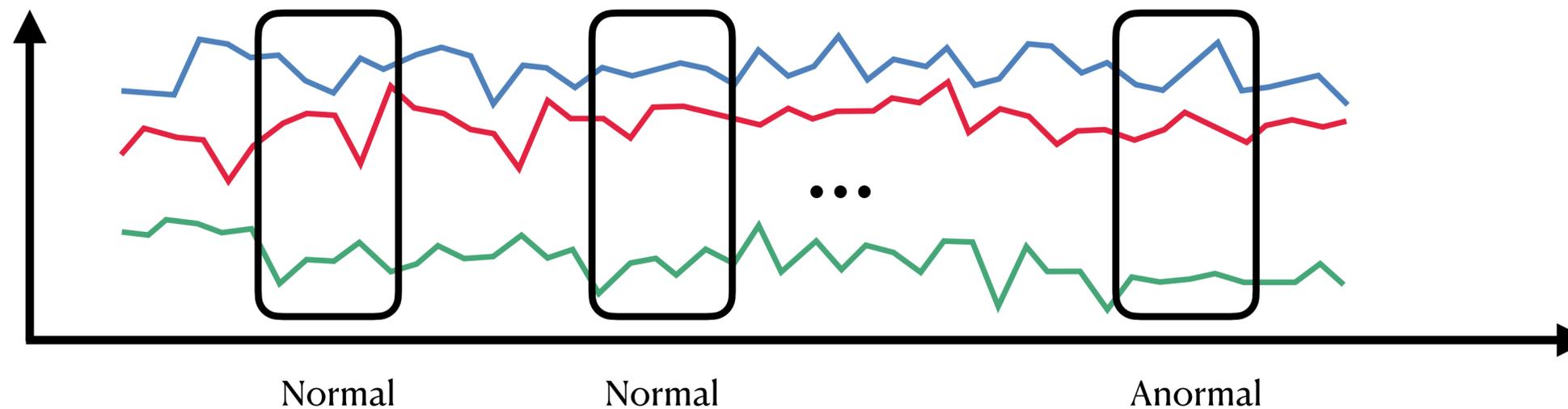
- A priori, pas de différence entre comportement normal et anormal



# Etape 5 : La preuve de concept    Data scientist

## Phase de R&D : présentation de la méthode

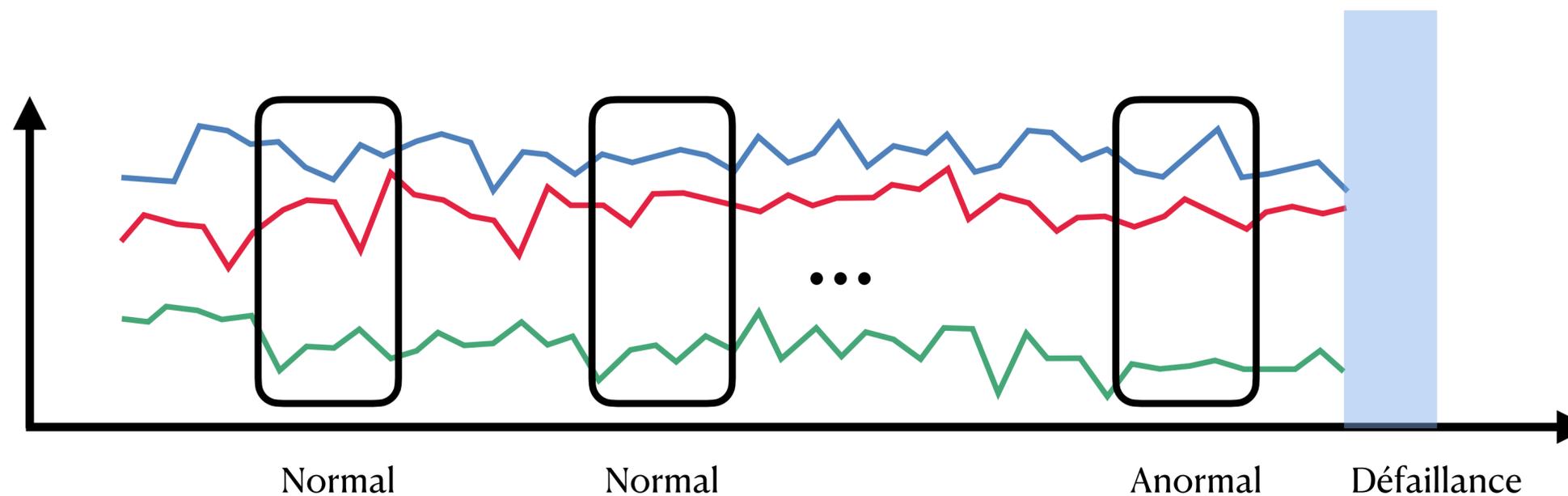
- A priori, pas de différence entre comportement normal et anormal



## Etape 5 : La preuve de concept    Data scientist

### Phase de R&D : présentation de la méthode

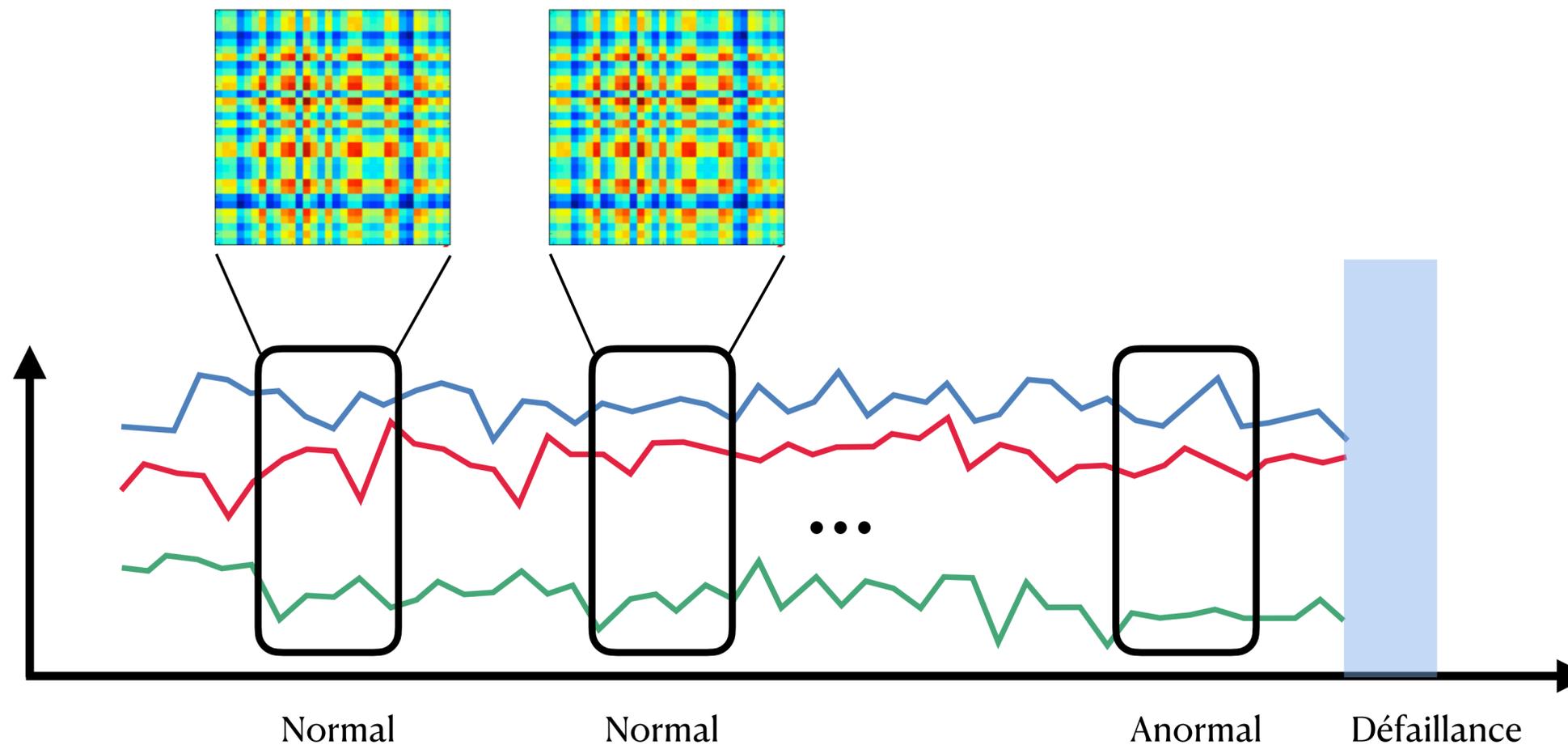
- A priori, pas de différence entre comportement normal et anormal



# Etape 5 : La preuve de concept    Data scientist

## Phase de R&D : présentation de la méthode

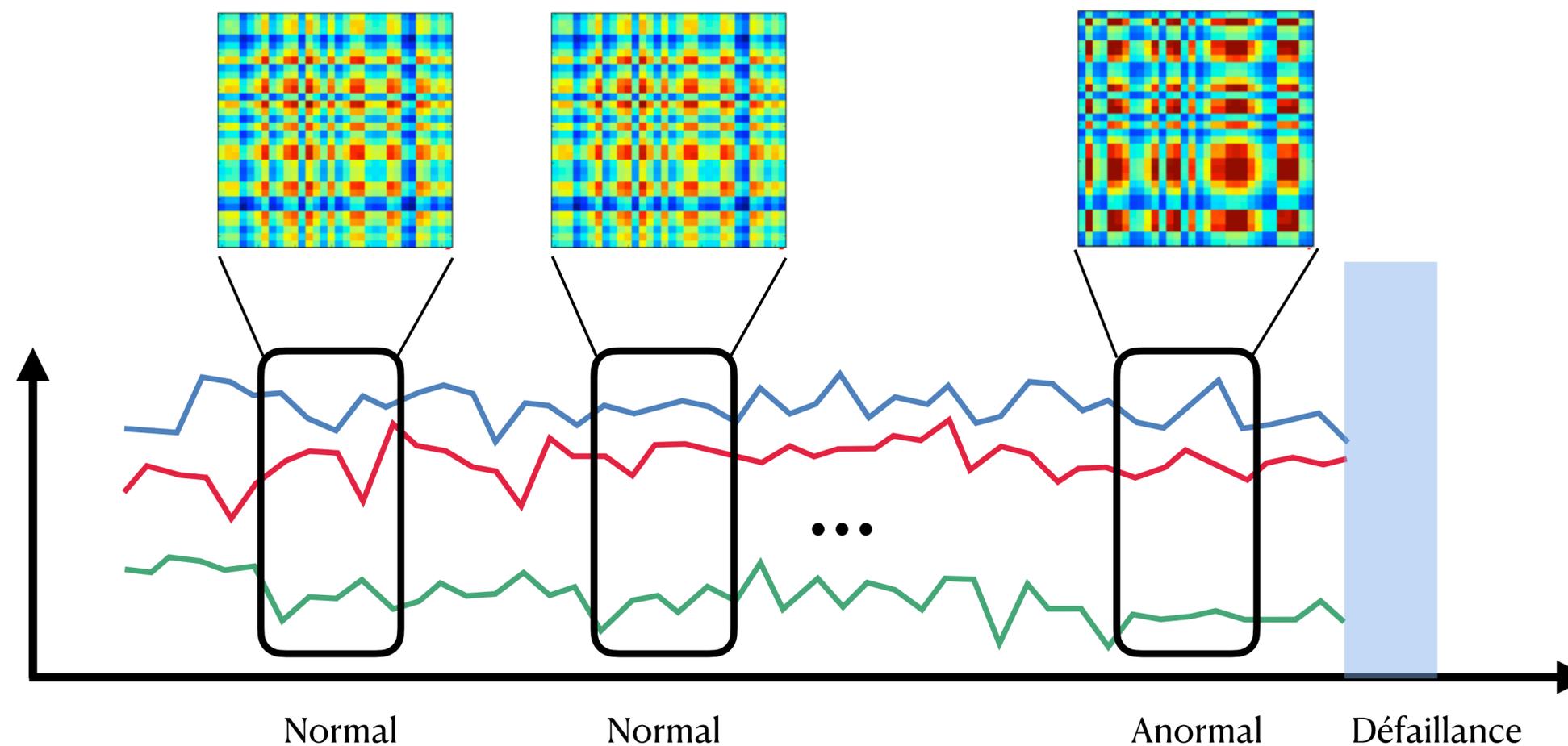
- A priori, pas de différence entre comportement normal et anormal
- Nouvelle observable : matrice combinant les paramètres initiaux



# Etape 5 : La preuve de concept Data scientist

## Phase de R&D : présentation de la méthode

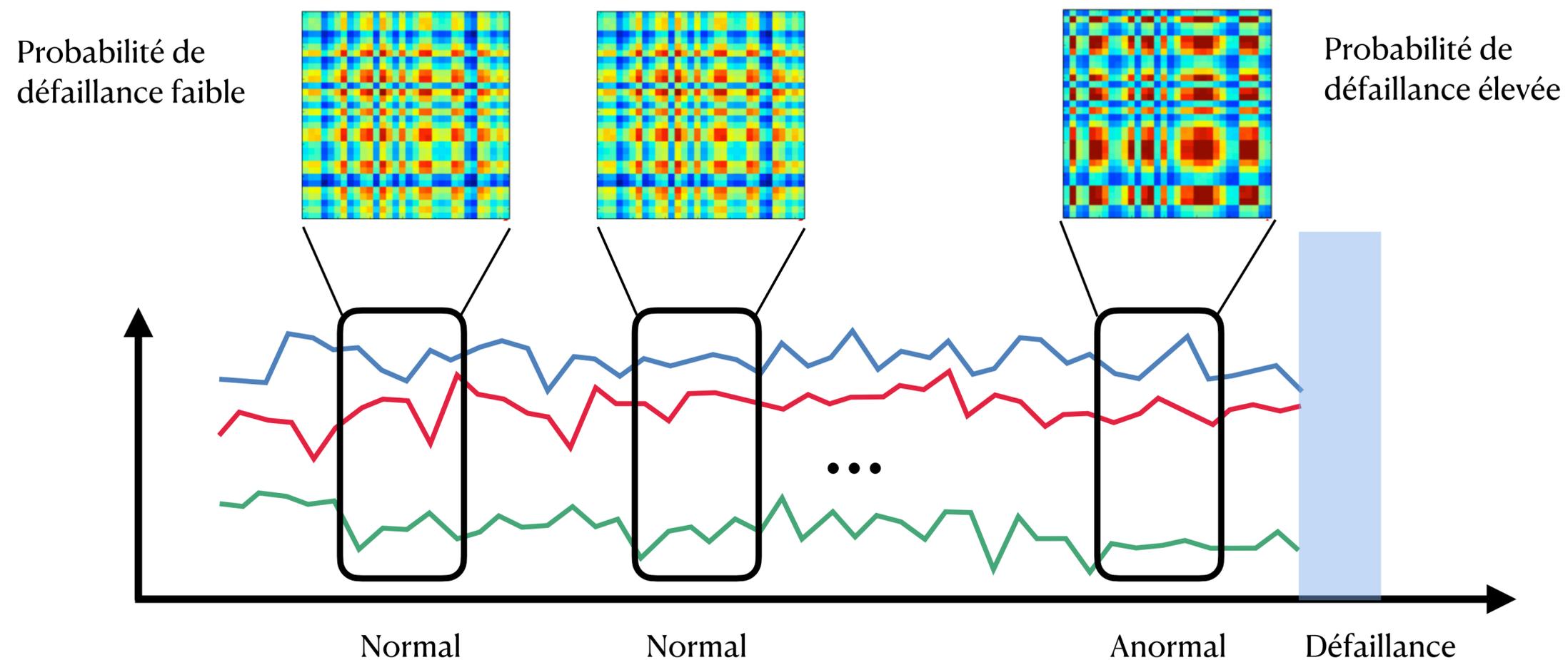
- A priori, pas de différence entre comportement normal et anormal
- Nouvelle observable : matrice combinant les paramètres initiaux
- Il trouve la combinaison de paramètres qui différencie les comportements



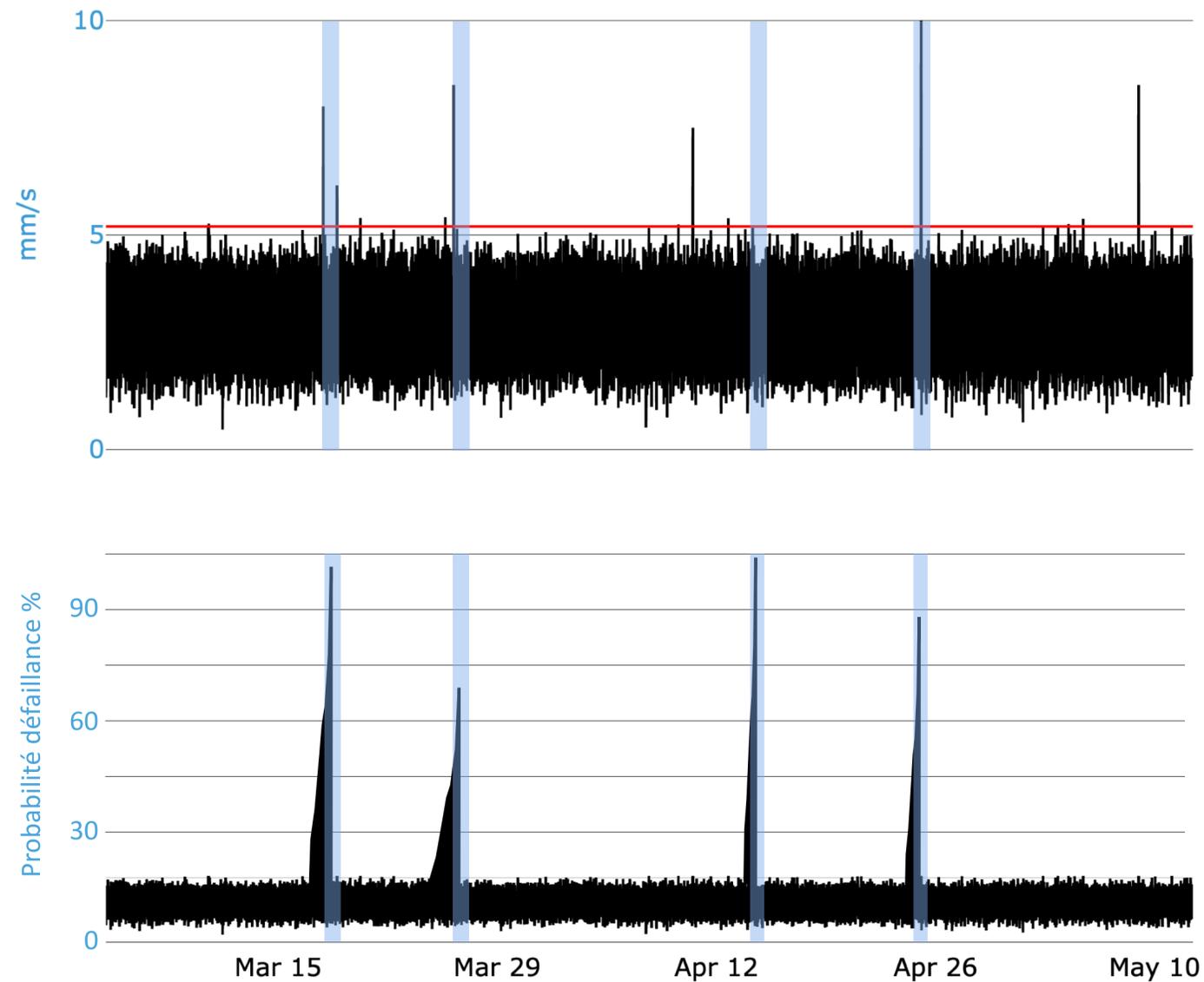
# Etape 5 : La preuve de concept Data scientist

## Phase de R&D : présentation de la méthode

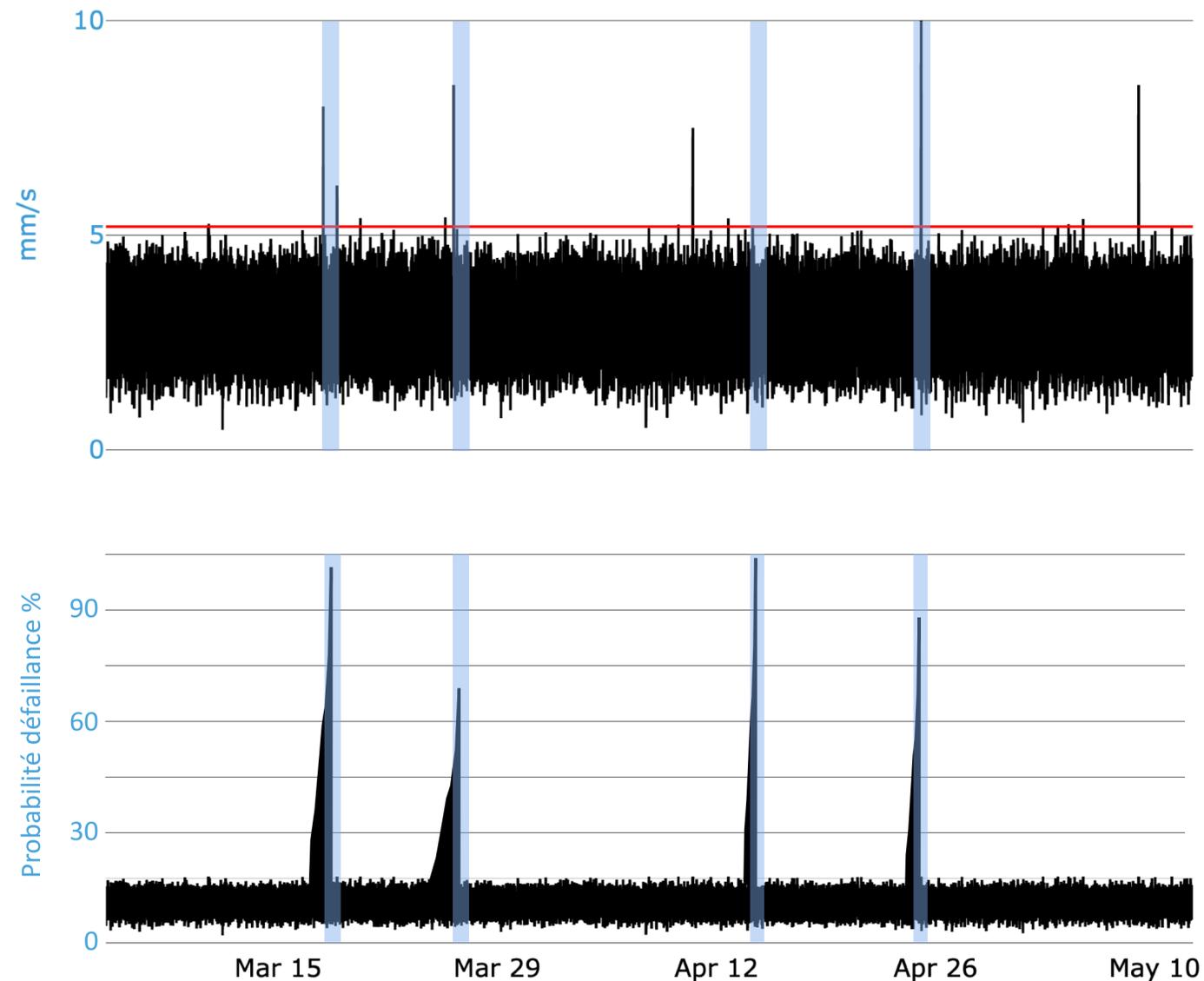
- A priori, pas de différence entre comportement normal et anormal
- Nouvelle observable : matrice combinant les paramètres initiaux
- Il trouve la combinaison de paramètres qui différencie les comportements



## Présentation des résultats

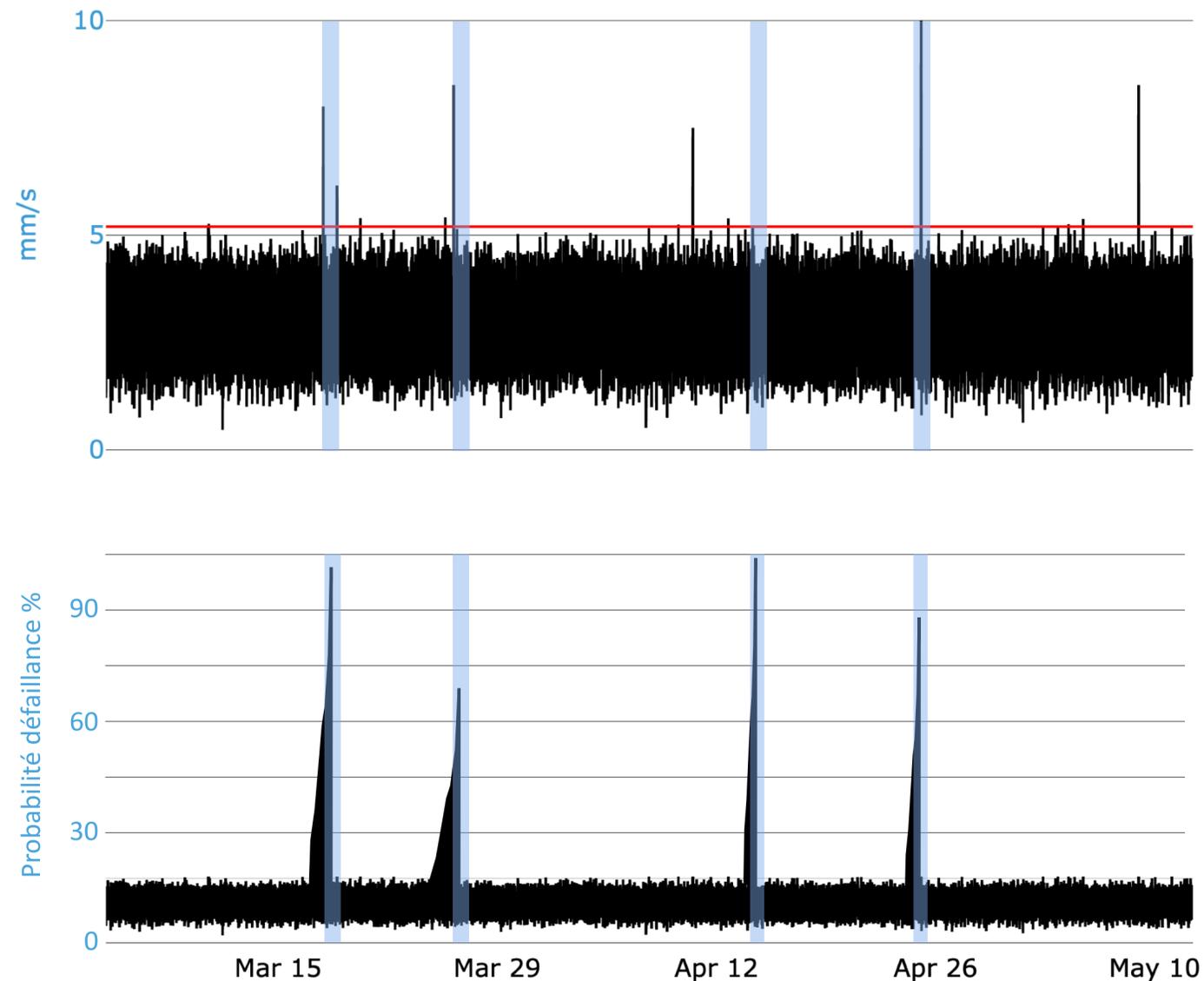


## Présentation des résultats



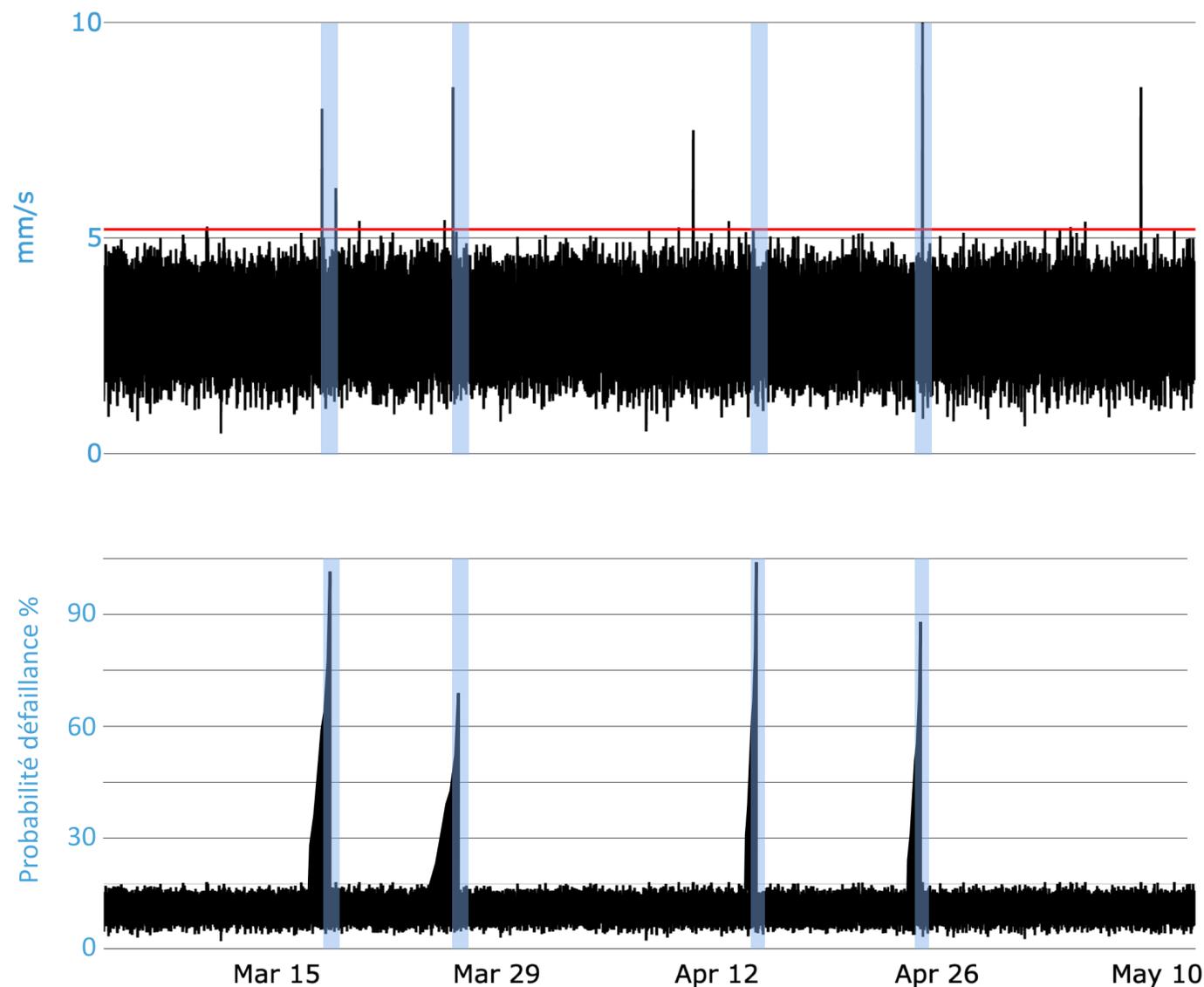
- Méthode moins sensible au bruit
- Pourcentage de fausses alertes diminué
- Pourcentage de pannes détectées amélioré

## Présentation des résultats



- Méthode moins sensible au bruit
- Pourcentage de fausses alertes diminué
- Pourcentage de pannes détectées amélioré
- Simulation du gain: **+23% de précision**

## Présentation des résultats



- Méthode moins sensible au bruit
- Pourcentage de fausses alertes diminué
- Pourcentage de pannes détectées amélioré
- Simulation du gain: **+23% de précision**

**Objectif business atteint**

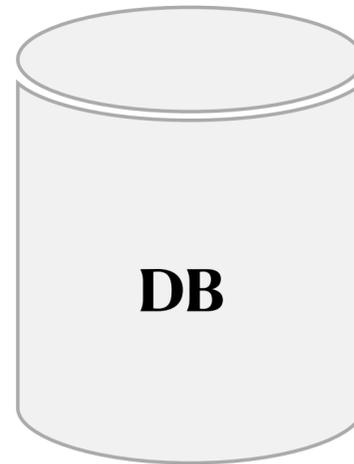
## Etape 5 : La preuve de concept

## Ingénieur informaticien

Road map industrialisation

# Etape 5 : La preuve de concept    Ingénieur informaticien

Road map industrialisation



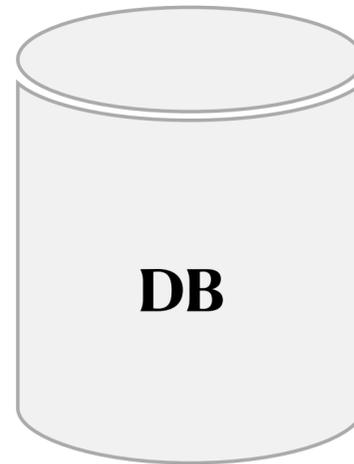
# Etape 5 : La preuve de concept    Ingénieur informaticien

Road map industrialisation



## Data

Production  
Raw Material  
Qualité  
Machines  
Autres?



# Etape 5 : La preuve de concept

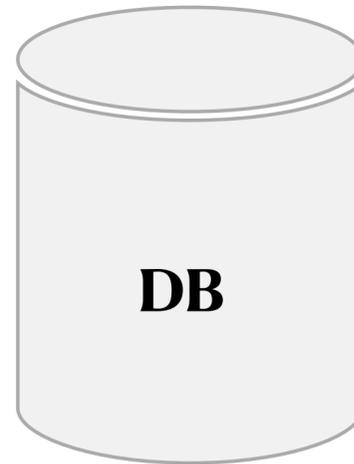
# Ingénieur informaticien

Road map industrialisation



## Data

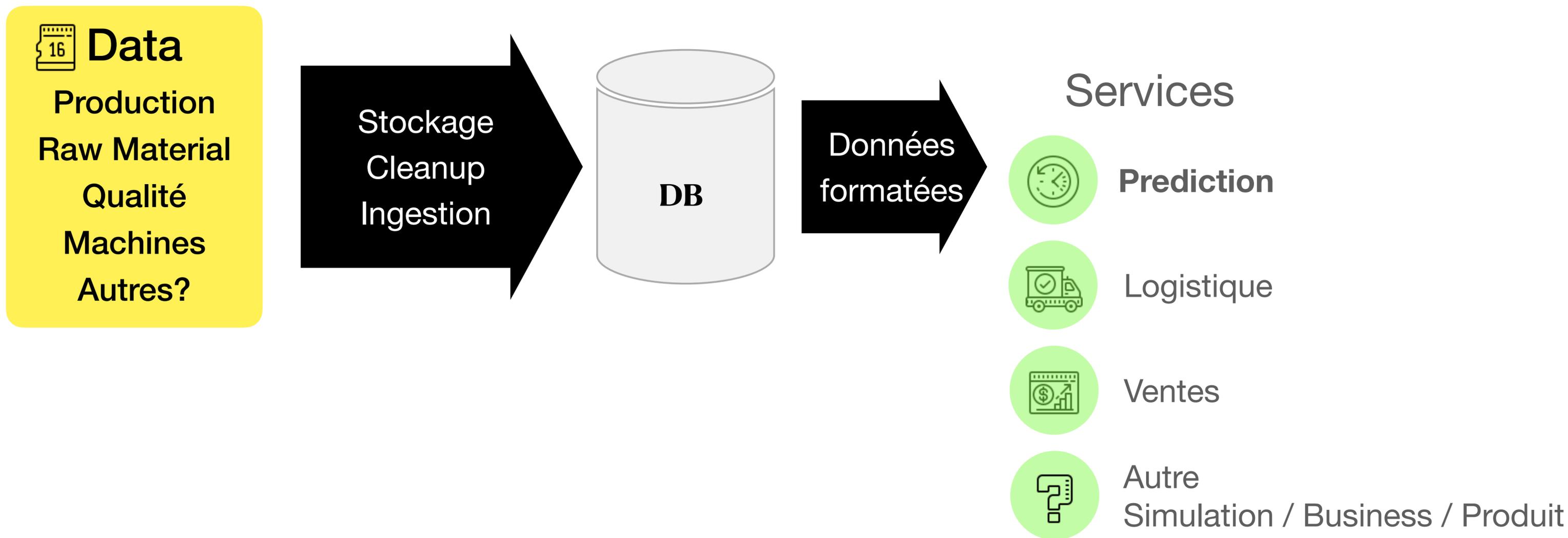
Production  
Raw Material  
Qualité  
Machines  
Autres?



# Etape 5 : La preuve de concept

# Ingénieur informaticien

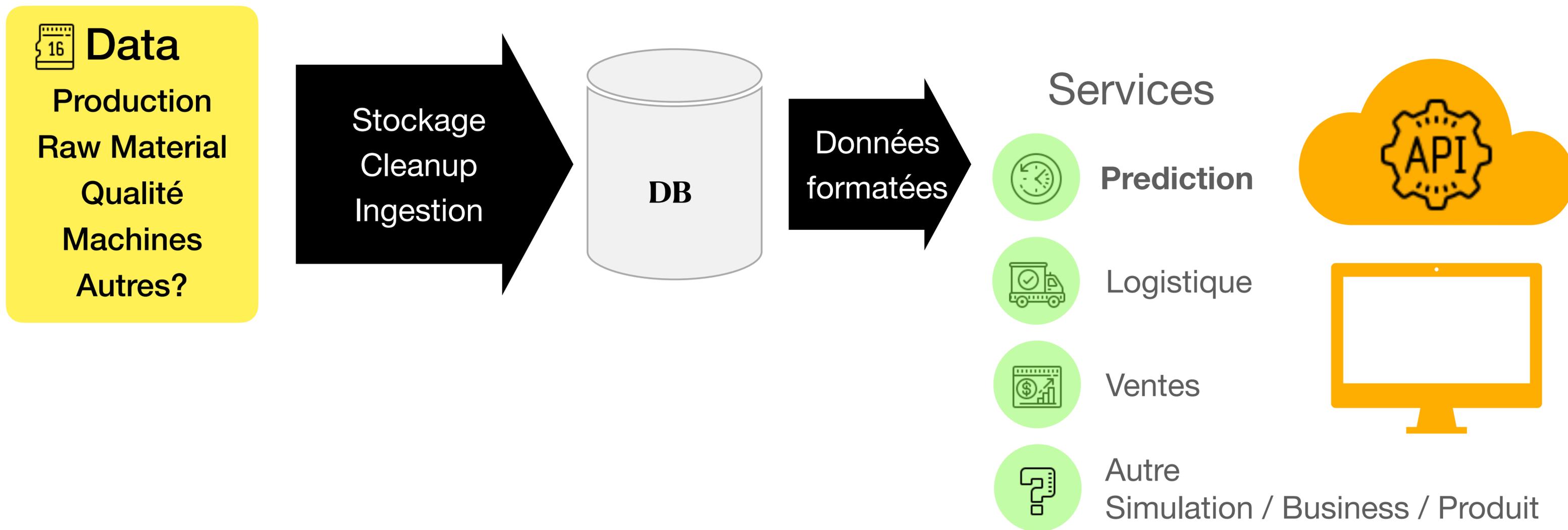
## Road map industrialisation



# Etape 5 : La preuve de concept

# Ingénieur informaticien

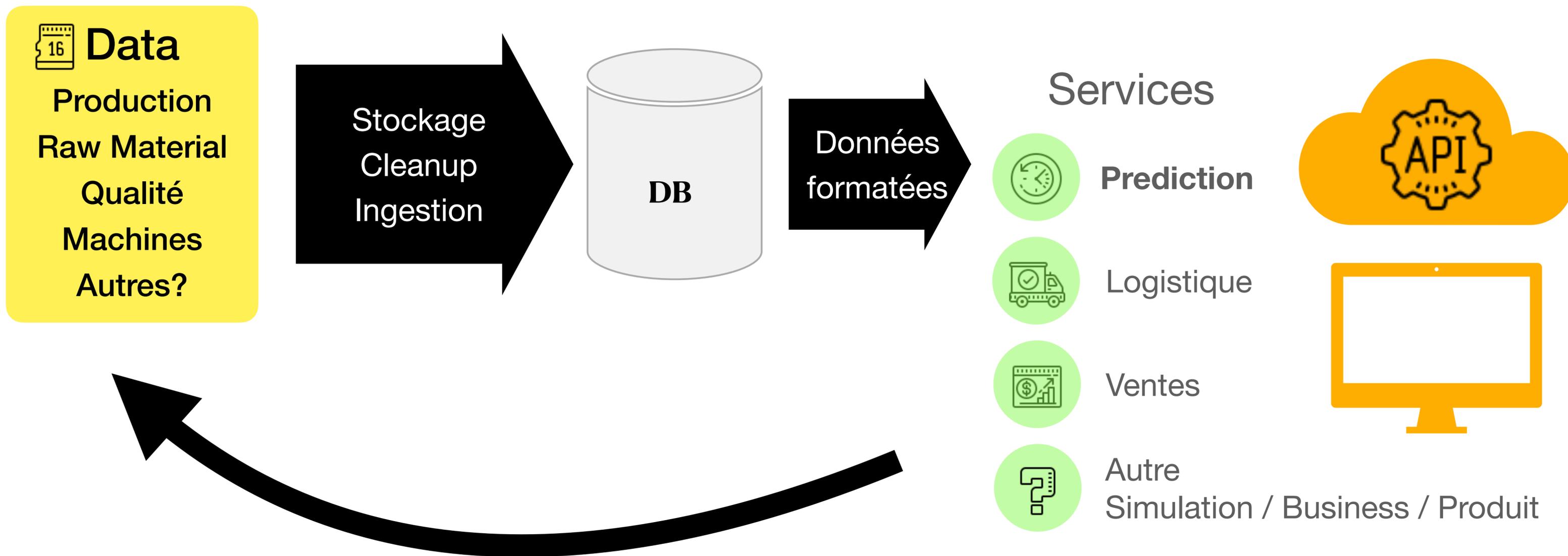
## Road map industrialisation



# Etape 5 : La preuve de concept

# Ingénieur informaticien

## Road map industrialisation



# Etape 5 : La preuve de concept

## Faisabilité

- ✓ • Traitement des données
- ✓ • Infrastructure
- ✓ • ROI

Validation du budget

## Etape 6 : Industrialisation

Suivi de la road map :

- Intégration de la solution dans l'infrastructure
- Automatisation
- Tests
- Retours d'expérience
- **Déploiement V1**

Fin du projet ?

## Etape 7 : Maintien de la solution

Pendant toute la durée d'utilisation :

- Stockage

- Sécurité

- Intégration

- Mises à jour



# Conclusions

Besoin

Solution

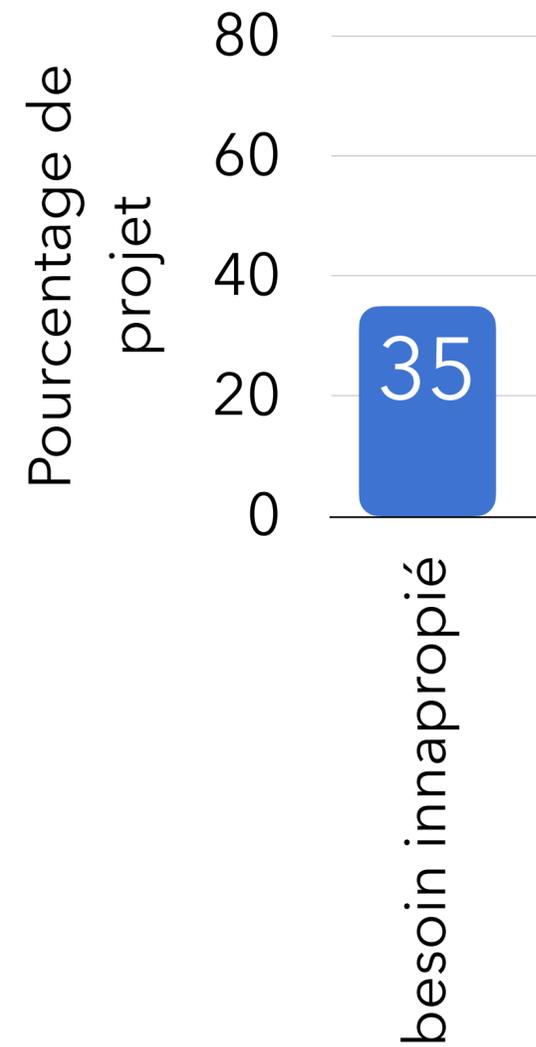
\*D. K. Becker, "Predicting outcomes for big data projects: Big Data Project Dynamics (BDPD): Research in progress," *2017 IEEE International Conference on Big Data (Big Data)*, 2017, pp. 2320-2330, doi: 10.1109/BigData.2017.8258186.



**IDEFICS**

# Conclusions

Besoin



Solution

Facteurs d'échec

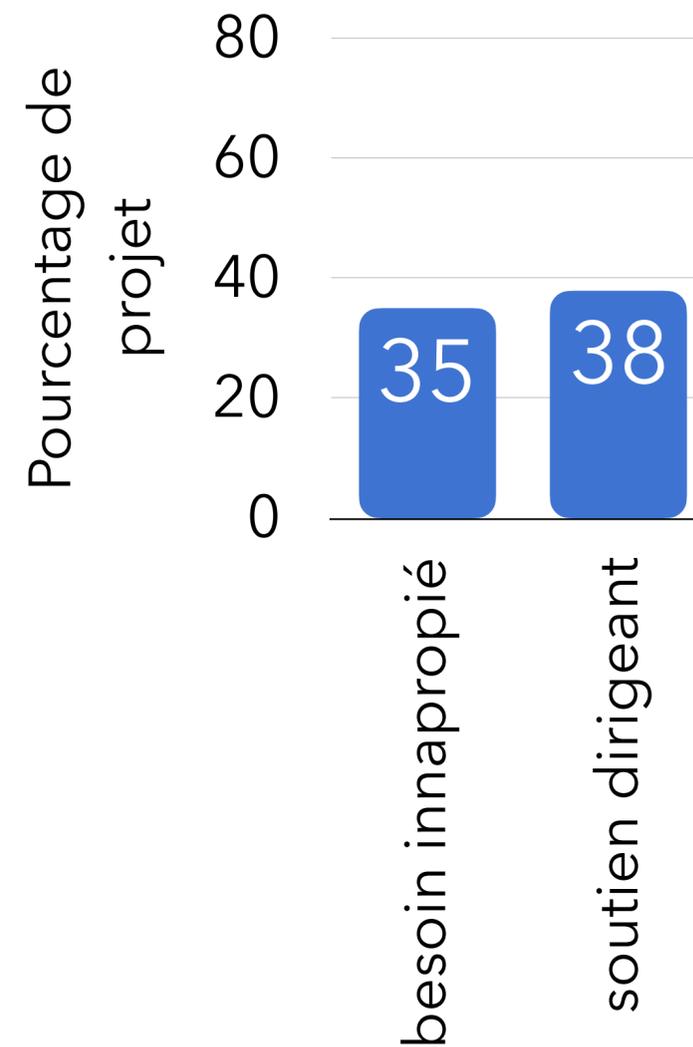
\*D. K. Becker, "Predicting outcomes for big data projects: Big Data Project Dynamics (BDPD): Research in progress," *2017 IEEE International Conference on Big Data (Big Data)*, 2017, pp. 2320-2330, doi: 10.1109/BigData.2017.8258186.

# Conclusions

Besoin

Convaincre

Solution



Facteurs d'échec

\*D. K. Becker, "Predicting outcomes for big data projects: Big Data Project Dynamics (BDPD): Research in progress," *2017 IEEE International Conference on Big Data (Big Data)*, 2017, pp. 2320-2330, doi: 10.1109/BigData.2017.8258186.

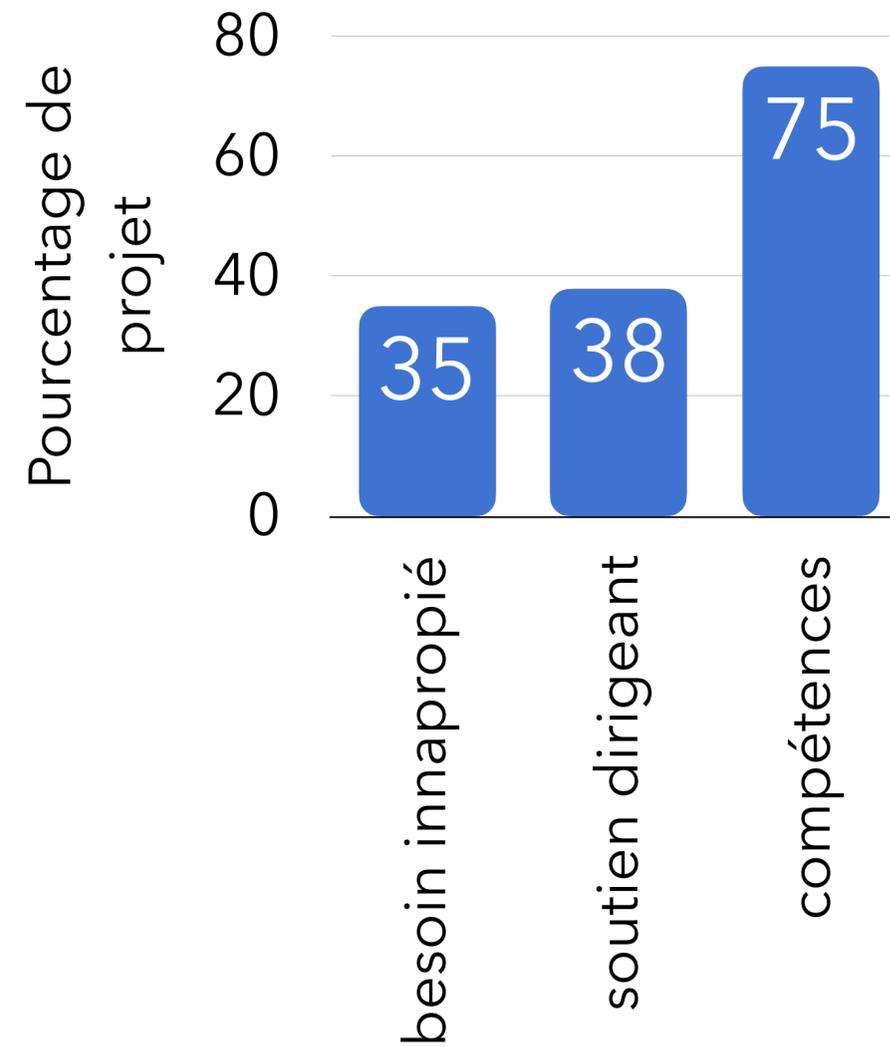
# Conclusions

Besoin

Convaincre

Equipe

Solution



Facteurs d'échec

\*D. K. Becker, "Predicting outcomes for big data projects: Big Data Project Dynamics (BDPD): Research in progress," *2017 IEEE International Conference on Big Data (Big Data)*, 2017, pp. 2320-2330, doi: 10.1109/BigData.2017.8258186.



IDEFICS

# Conclusions

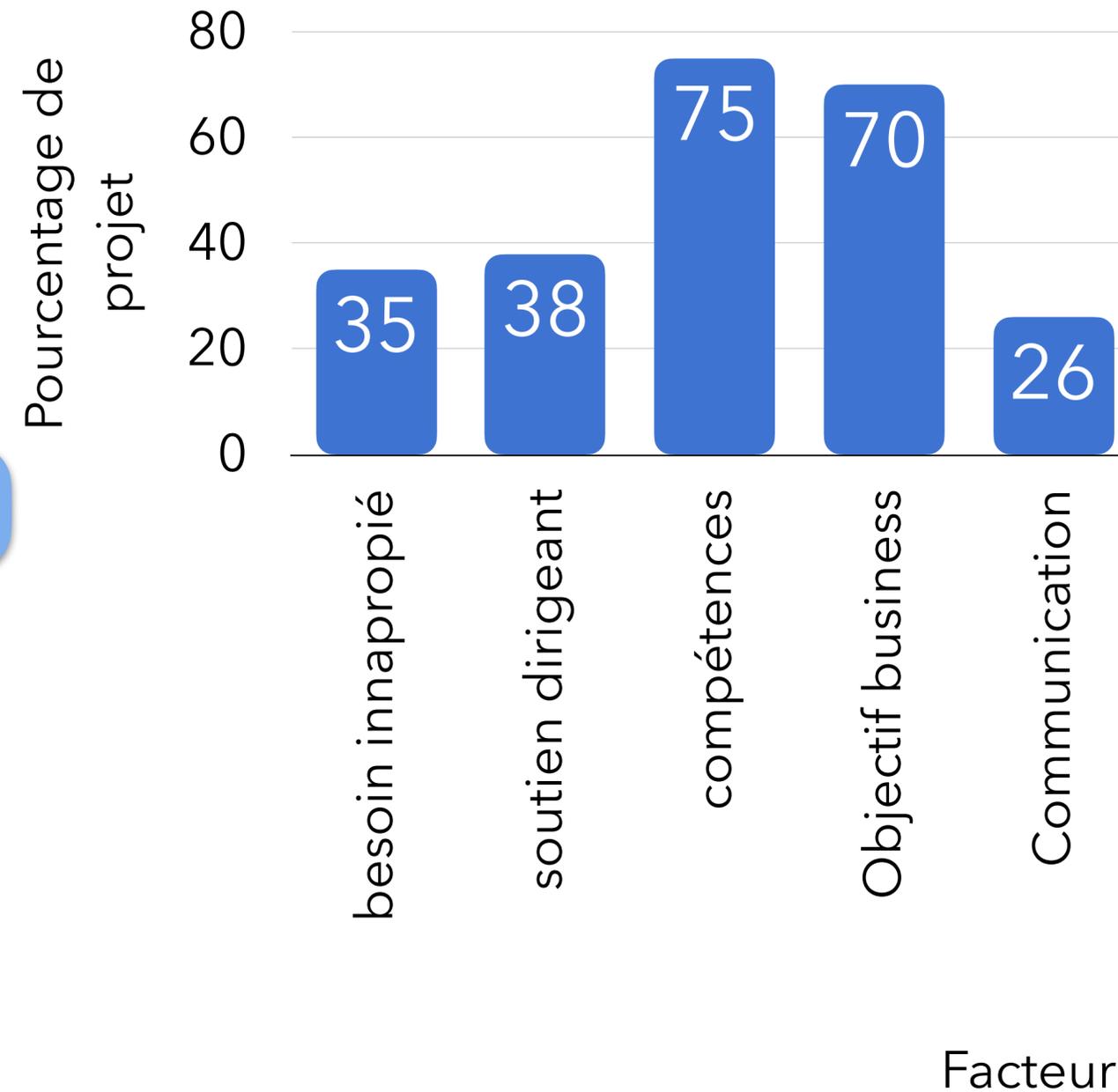
Besoin

Convaincre

Equipe

Définition du Projet

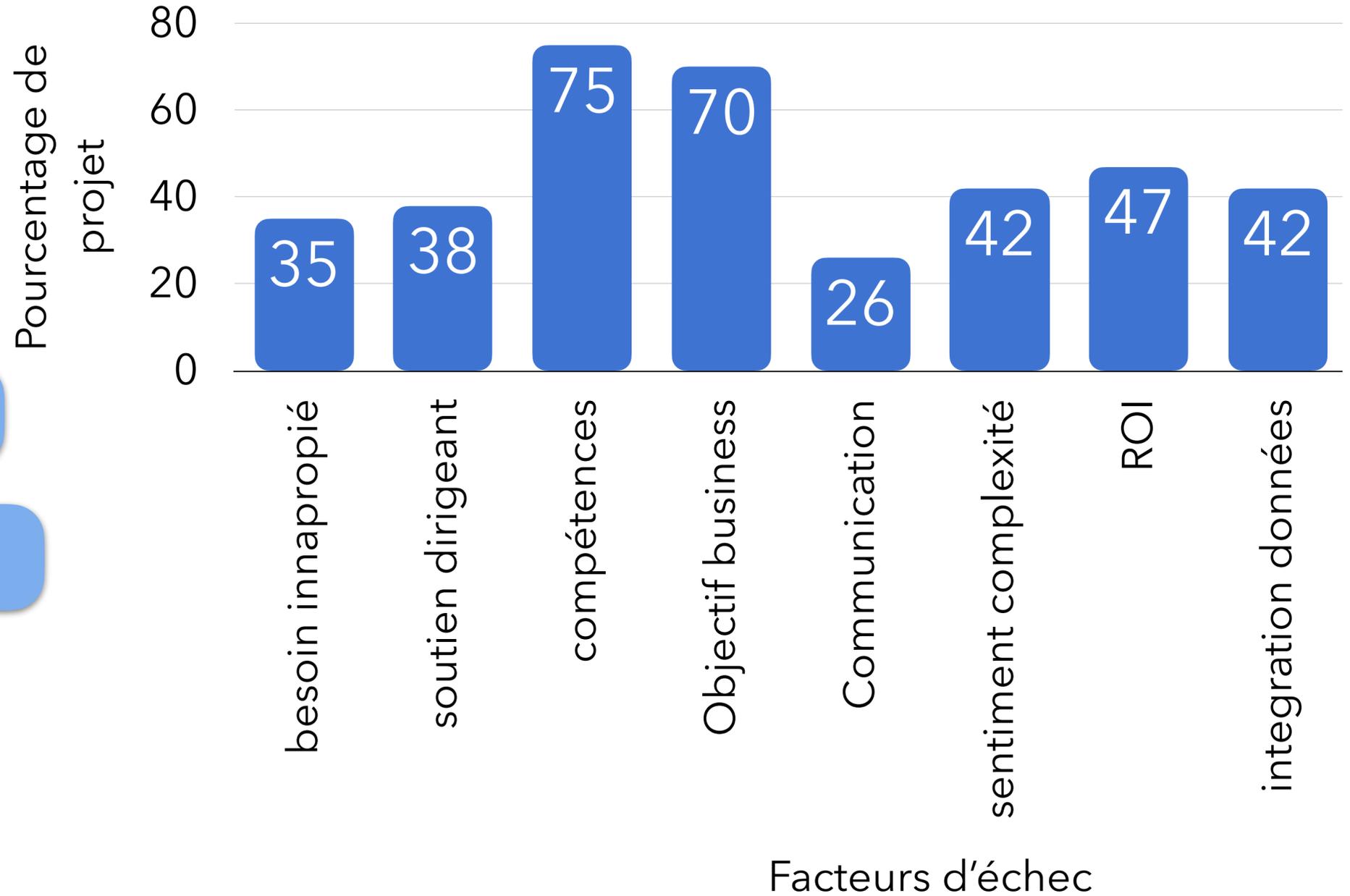
Solution



\*D. K. Becker, "Predicting outcomes for big data projects: Big Data Project Dynamics (BDPD): Research in progress," *2017 IEEE International Conference on Big Data (Big Data)*, 2017, pp. 2320-2330, doi: 10.1109/BigData.2017.8258186.

# Conclusions

- Besoin
- Convaincre
- Equipe
- Définition du Projet
- Preuves de concept
- Solution

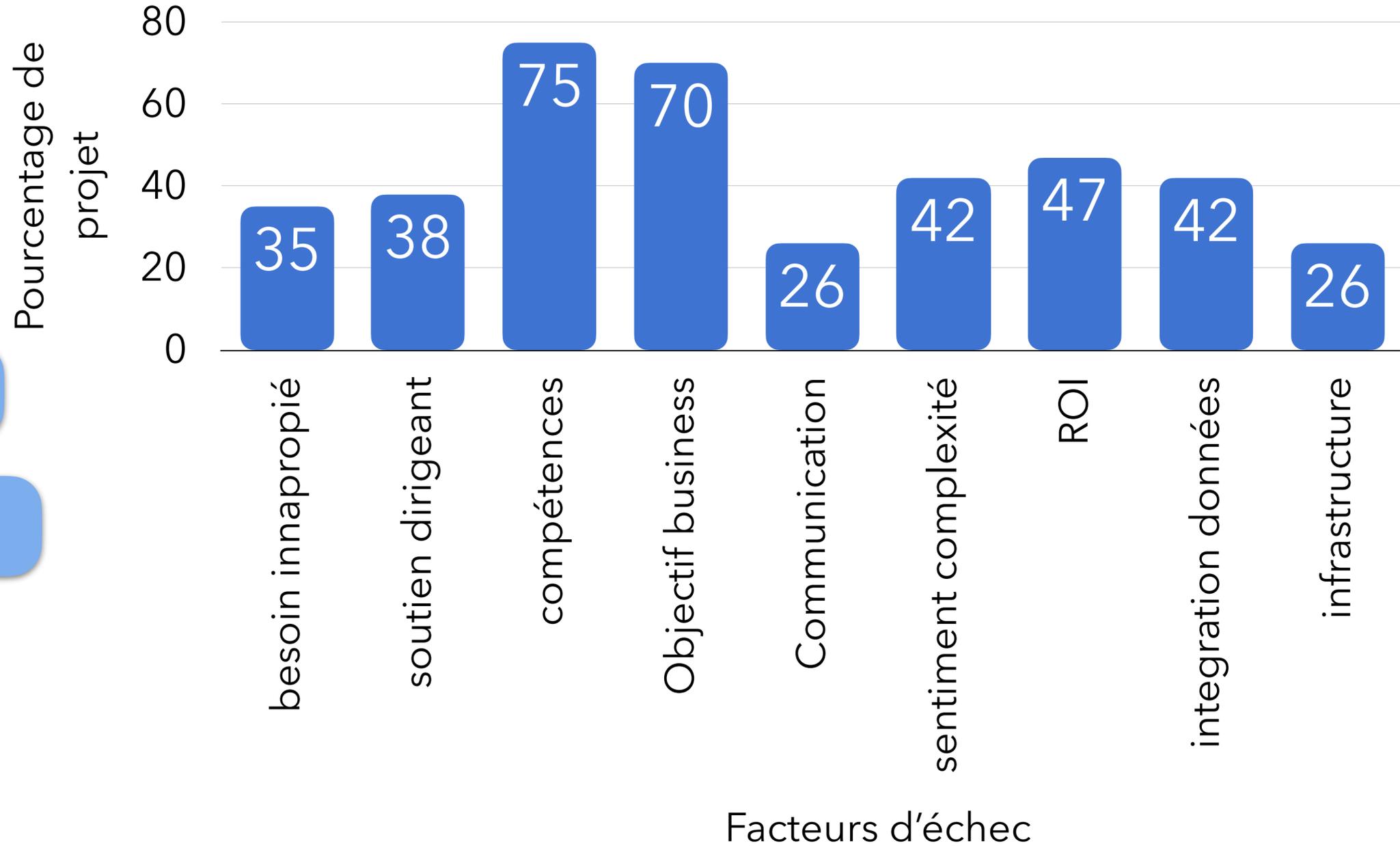


\*D. K. Becker, "Predicting outcomes for big data projects: Big Data Project Dynamics (BDPD): Research in progress," 2017 IEEE International Conference on Big Data (Big Data), 2017, pp. 2320-2330, doi: 10.1109/BigData.2017.8258186.



# Conclusions

- Besoin
- Convaincre
- Equipe
- Définition du Projet
- Preuves de concept
- Industrialisation
- Solution

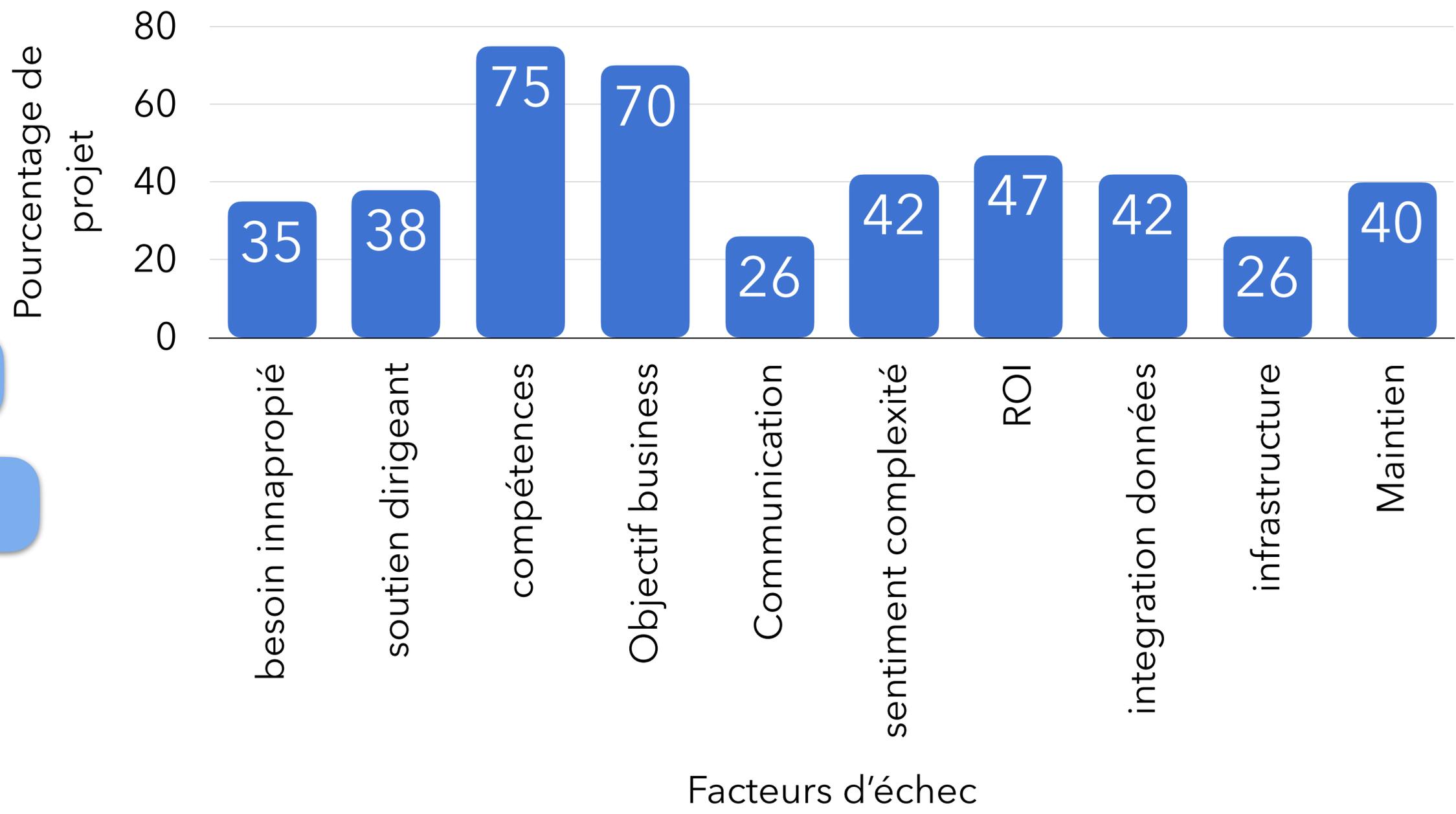


\*D. K. Becker, "Predicting outcomes for big data projects: Big Data Project Dynamics (BDPD): Research in progress," 2017 IEEE International Conference on Big Data (Big Data), 2017, pp. 2320-2330, doi: 10.1109/BigData.2017.8258186.



# Conclusions

- Besoin
- Convaincre
- Equipe
- Définition du Projet
- Preuves de concept
- Industrialisation
- Solution
- Maintien



\*D. K. Becker, "Predicting outcomes for big data projects: Big Data Project Dynamics (BDPD): Research in progress," 2017 IEEE International Conference on Big Data (Big Data), 2017, pp. 2320-2330, doi: 10.1109/BigData.2017.8258186.





# IDEFICS

Informatique, Données et Entreprises  
pour la Formation et l'Innovation en  
Calcul scientifique et pour la Société

<https://idefics.fr>

IDEFICS est cofinancé par la région Auvergne Rhône-Alpes et l'Union européenne dans le cadre du Fonds Européen de Développement Régional (FEDER).

