



Institut pour la **Maîtrise des Risques**
Sûreté de Fonctionnement - Management - Cindyniques

Big data in reliability

Projet de l'IMdR n° P17-4

Copyright IMdR – décembre 2019

Chef de Projet :
M. Vianney BORDEAU (RATP)

Contractants :
Sociétés LGM et QUANTMETRY



Quantmetry
Building AI with pioneers

L'Institut pour la Maîtrise des Risques tient à remercier :

- M. Vianney BORDEAU – RATP, Chef de projet, pour avoir dirigé cette étude,
- Les sociétés qui ont souscrit à ce projet et leurs collaborateurs qui ont participé à sa réalisation :

- | | | |
|---------------------|---|--|
| ➤ ALSTOM |  | MM. Pierre DERSIN et Stéphane GAUTHIER |
| ➤ DASSAULT AVIATION |  | M. Gilles DEBACHE |
| ➤ DGA |  | MM. Sylvain MEUNIER, Bruno LEBRETON et Lucien DEGARDIN |
| ➤ EDF |  | M. Emmanuel REMY |
| ➤ GRTgaz |  | Mme. Leïla MARLE |
| ➤ INRS |  | M. Pascal LAMY |
| ➤ IRSN |  | MM. Nicolas DECHY, Jean-Marie ROUSSEAU et Jean-Luc STEPHAN |
| ➤ NEXTER |  | M. Bruno COLIN |
| ➤ RATP |  | Mme. Sabrina IKHLEF et M. Vianney BORDEAU |
| ➤ SNCF |  | MM. Benoit GUYOT et Gilles PETITET |
| ➤ THALES |  | M. Victor LEYRONAS |
- La société LGM représentée par MM. François ESCUDIÉ, Christophe LE STIR et Frédéric DESCHAMPS ; la société QUANTMETRY représentée par MM. Martin LE LOC et Nicolas BOUSQUET,
 - M. André Lannoy animateur de la commission Produits de l'IMdR.

NOTE DE PRESENTATION ET DE SYNTHÈSE

Etat de l'art du Big Data pour la fiabilité

La généralisation de l'usage du numérique au sein des activités industrielles entraîne une **croissance et une variété importantes des sources de données à disposition des industriels**. La conjonction d'une **baisse du coût des technologies** nécessaires à leur traitement ainsi que de **nouvelles techniques** augmentant la capacité à les exploiter est une véritable opportunité à saisir pour **améliorer les stratégies et les méthodes de fiabilité**.

Afin de réussir cette transformation, il est important d'appréhender et de mettre en œuvre, en complément des pratiques actuelles, de nouvelles techniques disponibles (*Big Data* et techniques d'apprentissage machine notamment) dont l'état de l'art est développé dans cette étude.

Pour tirer pleinement parti de ces méthodes, l'amélioration de la fiabilité par le *Big Data* nécessite un **partage de l'information entre tous les acteurs de la fiabilité**, depuis la conception du produit jusqu'à son exploitation, en passant par sa fabrication. Afin d'y parvenir, nous recommandons d'engager en complément des actions individuelles une **stratégie de partage systématique de données** fournisseurs (tests de qualité), fabricants (protocoles de fabrication, mesure de la qualité en ligne, essais) et exploitants (usages, profils de fonctionnement, retours d'expérience).

Impact sur le REX et son analyse

Un des impacts majeurs de l'adoption du *Big Data* sur le retour d'expérience est de **transformer la connaissance du fonctionnement et du dysfonctionnement d'un système d'un mode statique à un mode dynamique et continu**. Les techniques de *Big Data* vont ainsi généraliser l'utilisation systématique d'un panel de données plus larges et plus fréquentes, voire l'utilisation de données jugées jusqu'alors inexploitable de par leur complexité à être rapprochées des données habituellement utilisées dans le cadre du retour d'expérience.

La **connaissance dynamique et continue des systèmes** va permettre à l'industriel ou à l'exploitant de **modéliser un jumeau numérique** de son système, pouvant aller jusqu'à un jumeau numérique **individualisé** de chacun de ses systèmes. L'objectif est de **passer progressivement d'une approche globale de REX à une approche personnalisée et individualisée**.

Le rapport du Projet de l'IMdR n°P15-2 "*Health and Usage Monitoring System (HUMS) – Health Monitoring*" (2017) concluait que les nouvelles technologies *HUMS PHM* génèrent une importante quantité de données et qu'un futur challenge à relever serait le nettoyage, le traitement et le stockage de ces données.

La **gouvernance de ces données** de capteurs et de l'ensemble des données utilisées dans le retour d'expérience est ainsi identifiée aujourd'hui **comme un prérequis indispensable à l'utilisation du *Big Data*** et de l'Intelligence Artificielle.

Il reste nécessaire de souligner les freins (ou contraintes) ou les bénéfiques (ou apports) que l'on peut rencontrer lors de la mise en place d'un projet de *Big Data*. Ainsi le **rôle de l'expert** dans une telle démarche **s'avère primordial**, que ce soit lors de la mise en place du projet (quelles données surveiller et collecter ? À quelles fréquences ?), ou lors de la phase d'analyse des données (labellisation du jeu de données d'entraînement par exemple). Ces nouveaux modes d'analyse nécessitent un **investissement en ressources humaines et financières** supérieur, tout du moins dans un premier temps, au niveau actuel requis par le REX.

Impact sur l'estimation de la fiabilité d'un composant, d'une fonction ou d'un système

Les techniques de *Big Data* vont aider les fiabilistes à **mieux appréhender le profil de fonctionnement des systèmes**, leur comportement vis-à-vis des agressions et ainsi à estimer plus précisément la fiabilité des systèmes qu'ils étudient. Les missions du fiabiliste vont s'en trouver modifiées, l'expert fiabiliste va de plus en plus être impliqué dans les métiers d'analyste de données (*Data scientist, Data analyst, responsable visualisation Data*) afin de définir au plus tôt quels équipements, quelles données, ou quelles évolutions il va falloir surveiller dès la mise en service.

En phase d'utilisation, **la connaissance dynamique que peut apporter le *Big Data* va permettre aux ingénieurs de maintenance de prévenir les défaillances au plus tôt** et ainsi de s'y préparer, soit en décalant l'occurrence de la panne en **modifiant le facteur de charge** ou d'autres paramètres influents, soit en **mettant en place le système de soutien adéquat** pour que tous les moyens pour l'intervention de maintenance soient disponibles à un instant optimisé. Ces prévisions vont donc modifier l'organisation de la maintenance d'un point de vue de la planification mais également avoir un **impact sur les compétences, les stocks, les moyens de détection...**

Impact sur les méthodes de maintenance (prévisionnelle)

L'impact du *Big Data* sur la fiabilité se diffuse sur l'ensemble des domaines en interaction avec ce métier. **La fiabilité et sa prévision sont la clé de voûte des maintenances** préventive et corrective. Elles permettent le dimensionnement du système de soutien.

En conception le *Big Data* et son volet IA permettent d'envisager une maintenance conditionnelle basée sur l'acquisition, le traitement et l'analyse d'un grand volume de données de fonctionnement du système précédent. Pour la mettre en place dès la conception, les outils classiques de conception de la maintenance préventive permettent d'évaluer les facteurs de risques du système, puis de sélectionner les points « critiques » à surveiller et à instrumenter. Le choix des capteurs doit faire l'objet d'une démarche rigoureuse, à mener par le fiabiliste et le bureau d'études, pour obtenir des données pertinentes et fiables : position et fiabilité des capteurs, risques de dérives, choix des seuils d'alerte, type d'algorithmes d'évaluation de la *RUL* (durée de vie restante)... Il faut aussi concevoir toute la chaîne de cette maintenance numérique ; plus on le fait tôt dans le processus, et plus on peut compter sur un outil de diagnostic du système plus performant, voire « intelligent », sur une maintenance prévisionnelle efficace, mais aussi sur un outil de REX enrichi avant même le début de l'utilisation du système. Tous ces éléments sont à intégrer, valider et qualifier dans le même temps que le système est pris en compte pour le développement de chaque élément du soutien logistique.

Lorsque le système est en utilisation, le *Big Data* permet de mettre en place la maintenance préventive conditionnelle ou prévisionnelle, dans le but d'intervenir avant la défaillance en fonction de seuils d'alerte, voire de pronostics sur la *RUL*, autant que possible en tenant compte de l'exploitation à venir. Le *Big Data* permet aussi de réaliser un REX rapide et enrichi, permettant de confirmer les facteurs de risques identifiés en conception, voire d'en identifier de nouveaux qui seront analysés par les experts, permettant ainsi d'ajuster la maintenance mise en place et d'identifier des évolutions. Deux points de vigilance sont cependant à prendre en compte et peuvent complexifier l'approche : l'augmentation des interventions préventives devrait diminuer le nombre de défaillances ce qui peut apporter une nouvelle incertitude sur la loi de comportement en fiabilité ; et par ailleurs les seuils définis pour cette maintenance peuvent évoluer dans le temps.

Impact sur l'aide à la décision

Face à l'utilisation croissante des initiatives *Big Data* au sein des entreprises, se posent naturellement la **question de la confiance dans les résultats** et de leur impact sur la prise de décision. Ces développements récents ne visent pas à remplacer les experts mais bien au contraire à les aider dans leur quotidien. Certains auteurs défendent désormais l'idée que l'IA actuelle n'est pas une Intelligence Artificielle mais bien une **Intelligence Augmentée** car elle reste mono tâche par construction et reste donc au service des experts métiers.

La confiance dans un algorithme repose avant tout sur une **coconstruction par l'expert et le *Data scientist*** pendant sa conception, que ce soit, à travers le choix d'une mesure de sa performance ou dans la sélection et la préparation des données en amont. Ces étapes sont primordiales pour assurer la qualité des résultats et donc bâtir un premier niveau de confiance.

Afin d'accélérer l'adoption des méthodes de *Big Data*, il convient également de souligner **l'importance de la communication** des résultats obtenus qui doivent être le plus **interprétables** possibles pour laisser la place à la critique, **mais aussi correctement représentés** afin d'accélérer la prise de décision.

La confiance dans les algorithmes de *Big Data*, qu'ils soient « boîte noire » ou non, doit reposer sur une **reproductibilité des résultats obtenus**, des **validations successives** de la part de l'expert lors de leur conception et un **suivi régulier de leur performance** une fois en production. L'IA possède le potentiel

d'être plus performante que les outils traditionnels d'analyse statistique à condition d'avoir suffisamment de recul dès l'amont et un soutien continu des experts métiers.

Impact sur l'organisation et les compétences

Le *Big Data* et les techniques associées ont un **impact significatif sur les organisations et les compétences au sein des entreprises**. En effet, la réussite de ce type d'initiatives passe par une cohérence d'ensemble entre les objectifs métiers (fiabilité, maintenance, ...), les ressources employées (compétences, données, environnement technique) et les méthodes employées pour piloter les projets *Big Data*. Afin d'orchestrer cette transformation, **les entreprises se dotent de modèles d'organisation dédiés à l'exploitation de la donnée** en fonction de leurs spécificités (d'un modèle centralisé à un modèle totalement externalisé).

Un autre impact est la nécessité d'accorder les temporalités en cycle produit et cycle projet IA. Ces projets n'ayant pas le même horizon temporel, le **succès d'un projet IA repose avant tout sur la création d'un binôme expert - *Data scientist*** qui permettra de garantir une cohérence d'ensemble entre ces deux expertises.

Enfin, la valorisation des données nécessitent **des compétences et des responsabilités nouvelles** qui peuvent être développées par **des moyens de formation variés** (*Bootcamp, MOOCs, meetups, ...*).

Traitement d'un cas test

Un cas test a été réalisé afin d'illustrer les différentes tâches du projet. L'objet est la création d'un outil d'aide à la décision destiné à évaluer la probabilité d'une pièce soupçonnée défectueuse de l'être réellement, de façon à éviter la non-détection de défaut lors du renvoi au fournisseur (NFF). Un jeu de données a été fourni par Dassault Aviation. Cette aide à la décision nécessite la conception et la **mise en œuvre d'un algorithme supervisé de classification qui évalue la probabilité de NFF** en fonction d'un certain nombre de critères issus de données connues (trames électroniques de l'avion, ...). Certaines données sont structurées (données de description, données issues de capteurs...), d'autres non structurées (texte libre sur fiches scannées, comptes rendus de réparation...). L'étude s'est concentrée sur la proposition de solutions pour répondre aux **deux contraintes majeures** dans la préparation des données :

- **Exploitation de données non structurées** : extraction d'informations depuis les **comptes rendus de réparation "authorised release certificate"**,
- **Rapprochement d'événements sans clés naturelles d'association** : création de continuité numérique pour identifier et rapprocher les événements décrivant une défaillance et l'envoi en réparation de la pièce associée.

Fourniture d'un guide de bonnes pratiques

Sur la base des travaux et réflexions du projet, un guide de bonnes pratiques a été élaboré à destination de tout industriel souhaitant se renseigner, spécifier ou mettre en place un projet *Big data* appliqué à la sûreté de fonctionnement ou à la maintenance.

Perspectives

Le projet IMdR P17-4 peut se prolonger au travers de plusieurs voies d'étude ou recherche :

- Etude du développement d'une méthode pour systématiser une étude de maintenance prévisionnelle en conception.
- Comment équiper en capteurs un système existant, quelles phases respecter, quelles exigences de fiabilité des capteurs ?
- Estimer et comprendre les différences et points communs entre approche de fiabilité traditionnelle et approche *Big Data / IA* pour les études de fiabilité prévisionnelle.
- Estimer l'impact de l'utilisation de la méthodologie *Big data* sur le binôme expert - *Data scientist* : analyse de la situation de travail, impact du changement, rôle de chacun, efficacité.
- Assurer une veille sur la cartographie des normes et la standardisation en lien avec le *Big Data* ou l'IA et les applications pour la sûreté de fonctionnement.