



CGI BUSINESS
CONSULTING

N°12
Juillet 2022

CHROMATIQUE

Data

*Comment faire de la Data un
avantage comparatif ?*

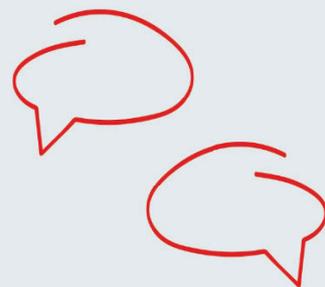




SOMMAIRE

Editorial	03
Data et principes de bonne conduite, immersion dans l'univers des jeux	05
Saisir l'avantage compétitif : les données comme levier d'optimisation pour la gestion du risque opérationnel	08
Séminaire annuel de l'AFCM : table ronde autour de la digitalisation du Post-Trade et article associé	17
Opportunités et défis des use cases Data pour le secteur financier	19
Comment tirer un avantage comparatif de la Data (appliquée aux stress tests climatiques)	22

EDITORIAL



Christophe Da Cunha et Franck Benzoni
Associés du Groupe Harwell Management

*C'est avec le plus grand plaisir que nous vous livrons ce mois-ci, l'ensemble des réflexions trimestrielles du **Lab Data** autour de la thématique « Comment faire de la data un avantage comparatif ? ». Ce thème, ayant fait l'objet d'une récente publication issue de la Harvard Business Review a inspiré chez nous cette idée de recueillir de manière formalisée, nos investigations et conclusions sur ce même sujet.*

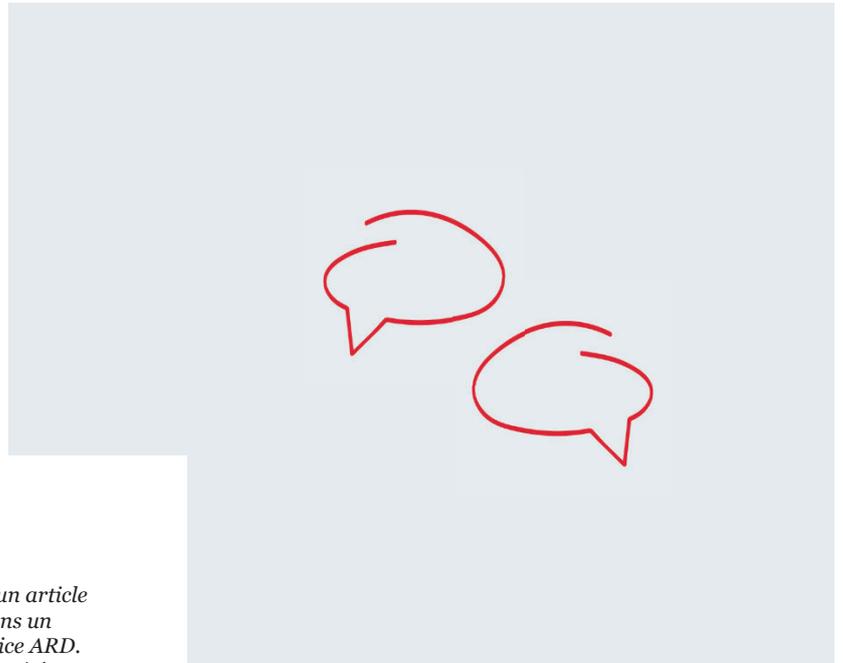
***Pascal de Lima**, chef économiste d'Harwell Management est l'auteur du premier article introductif « Comment tirer un avantage comparatif de la data ? ». Le terme « comparatif » est ici primordial. Pourquoi ? Parce que toute stratégie n'a de valeur que si elle est établie par rapport à un avantage que nous possédons sur nos clients. Cet article nous explique par ailleurs que la data ne peut être comparée au pétrole des entreprises, car à la différence du pétrole, celle-ci reste très peu exploitée par le secteur bancaire et financier. Tout du moins pas assez en tant qu'avantage comparatif.*

L'article fait également état des principes de bonne conduite ex ante avec cette idée que l'avantage comparatif de la data viendra aussi de la bonne utilisation de l'intelligence artificielle.

*Le second article s'intitule « Comment saisir l'avantage compétitif : les données comme levier d'optimisation pour la gestion du risque opérationnel ». **Hector Hernandez** et **Gabriele de Stefano** (Practice Conformité) vous proposent ici une analyse portant sur le bon modèle de données à sélectionner en insistant sur les techniques traditionnelles par opposition à l'irruption des techniques analytiques avancées face aux exigences réglementaires.*

S'ensuit un récapitulatif de la conférence animée par notre département BFI à l'AFCM (Arab Federation of Capital Markets) où Patrick Fosset a eu l'occasion d'aborder un sujet clé « Le rôle de la data dans la digitalisation du post trade ». Cette intervention a permis de faire le point sur la situation actuelle de la digitalisation dans la BFI et de revenir sur les éléments essentiels : motivations, sens des use cases, et facteurs de réussite.

*Nous remercions **Patrick Fosset** pour cette conférence réussie. Le lien de la vidéo est disponible au sein de la partie dédiée aux travaux de BFI. Fort de cette conférence, les équipes de BFI se sont focalisées sur la data et nous livrent, au travers de leur analyse, que dans toute initiative d'automatisation, la donnée est un composant primordial, surtout si les processus impliquent plusieurs intervenants sur les marchés. Nous remercions **Patrick Fosset** et **Aurélié Demangeon** pour leur contribution à cette chromatique.*



Ensuite, nous vous proposons de découvrir un article (également publié dans la Revue Banque dans un format « article de presse »), de notre practice ARD. Ithier de la Salle vous présentera les opportunités et défis des use cases data pour le retail banking en insistant plus particulièrement sur une méthodologie éprouvée qui permet, en acceptant les erreurs de laboratoire, de choisir la bonne fonctionnalité avec un ROI Garanti.

Le dernier article provient de David Ciolfi et Cyrille Almenar, de la practice TFFR « Comment tirer un avantage comparatif de la donnée » : application Stress Test avec en particulier un décryptage des données demandées dans le stress test.

Nous espérons que cette chromatique permettra, au-delà d'une prise de conscience du rôle de la donnée, de vous livrer l'intérêt d'un avantage comparatif pour lui donner du sens. Les enjeux climatiques et environnementaux sont essentiels et constituent, il est certain, un axe tangible dans la notion de « donner du sens » à la donnée.

Le secteur bancaire est de tout temps habitué à modéliser un nombre important de data pour des raisons de gestion des risques et de maximisation de la rentabilité.

Face à l'ampleur des enjeux climatiques, le climat s'est donc invité dans la gestion des risques globaux bancaires, ce qui a généré une prise de conscience des banques et leur responsabilisation. Cela étant, comme pour les risques traditionnels, la modélisation du risque climatique n'échappe pas à certaines limites et bien que le climat soit peut-être plus tangible que le cours d'une action, certaines méthodes contiennent des limites. Il est probable qu'à l'avenir, les autorités ou les banques elles-mêmes proposeront de les réduire via de nouvelles méthodes...

C. Da Cunha
F. Benzoni
Associés du Groupe Harwell



COMMENT TIRER UN AVANTAGE COMPARATIF DE LA DATA ? DATA ET PRINCIPES DE BONNE CONDUITE, IMMERSION DANS L'UNIVERS DES JEUX

On observe, en particulier dans les revues spécialisées et par ailleurs dans la presse, une très forte recrudescence du nombre d'articles consacrés à la data et la manière de faire des données, un avantage comparatif spécifique à l'entreprise. Quelle est l'ampleur de l'augmentation du volume de données ? Peut-on la comparer au pétrole ? En clair, la data est-elle « le pétrole des entreprises » ? Ensuite nous verrons que pour valoriser la donnée, il faut en réalité trois choses que l'on ne retrouve pas dans l'image du pétrole : (1.) des bases de données de qualité, (2.) des principes de bonne conduite, et (3.) surtout des innovations qui permettent de les exploiter comme l'intelligence artificielle.



1. La data est souvent désignée comme le pétrole des entreprises mais revêt pourtant des aspects très différents

Aujourd'hui selon IDC (Intelligence Data Corporation), la quantité de données générées chaque année dans le monde devrait passer à 175 ZO en 2025 via le développement du multi-connecté et donc, des capteurs. C'est l'internet des objets qui est responsable de cela. Pour donner une idée (itpro.fr) un seul zettaoctet représente plus de 40 000 milliards de DVD récoltés sur des supports variés d'objets connectés. Chaque jour,



nous interagissons 4800 fois avec un objet connecté soit une fois toutes les 18 secondes avec l'essor de l'IoT. La data n'est donc pas le pétrole des entreprises pour une raison simple : depuis la fin des années 1980 en variation, la consommation de pétrole est supérieure aux découvertes alors que l'utilisation elle-même par des outils de la data (sa consommation) est toujours inférieure aujourd'hui aux tétra données de data dont le potentiel est encore infini aujourd'hui. Comment alors tirer un avantage comparatif de la data face aux tétra données[1] ?

2. Des principes de bonnes conduites ex ante

→ ***Le premier principe** est le fait de comprendre que les données doivent servir un use case cible, une fonctionnalité cible qui correspond à un réel besoin sur le marché ou dans les organisations. On parle de **la connaissance de la valeur de la donnée**.

Mobileye, fournisseur de systèmes avancés d'aide à la conduite automatique propose des systèmes anti-collision et des alertes de franchissement de ligne blanche pour véhicules. Ici, ses clients sont des constructeurs automobiles qui réalisent eux-mêmes des tests. Mais en amont du système, c'est-à-dire en intrant du système, ce sont les données de ces tests qui ont de la valeur et qui permettent d'atteindre des niveaux de fiabilité de 99,9% grâce à l'outil ex post.

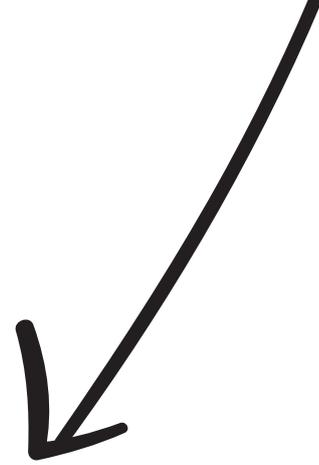
→ ***Le second** est cette idée qu'il n'est pas toujours nécessaire d'accumuler une quantité délirante de données. Là aussi, qualité n'est pas quantité. Il y a un moment donné **où la valeur marginale de la donnée risque de décroître**. Il faut donc une quantité optimale de données constamment renouvelées et de telle sorte que sa valeur marginale ne décroisse pas.

Zynga, qui a lancé « Farmville » aborde ce fait. En effet, leurs données ne sont pas devenues obsolètes d'une semaine à l'autre. Après avoir lancé d'autres jeux à succès dont « **Farmville 2** » et « **Cityville** » cette société de jeux sociaux a stoppé sa production jusqu'à perdre plus de la moitié de sa base d'utilisateurs. Ensuite, elle a été supplantée par **Supercell** (Clash of Clans et Epic Games (Fortnite)). Après avoir atteint un pic en 2012 à 10,4 milliards de dollars, la valorisation de Zynga a ensuite été divisée par plus de deux les six années suivantes. Un bel exemple d'obsolescence des données en intrant des jeux produits.

→ ***Le troisième** est que les données doivent être exclusives à l'entreprise et ne **doivent pas être imitables**.

Adaviv est une startup de la région de Boston. Elle propose un système de gestion des récoltes pour surveiller chaque plan en continu. On utilise l'IA avec des logiques de vision, et l'on suit des indicateurs de plans invisibles à l'oeil nu par la biométrie. L'intelligence artificielle permet dans le cas présent, de produire des signaux avant-coureurs de maladies et carences en nutriments qui amènent à des rendements décroissants. Mais en intrant, ce sont des données non imitables et exclusives à l'entreprise des utilisateurs. Plus le **nombre** d'utilisateurs Adaviv augmentait, plus l'outil pouvait recueillir des informations sur les variations de données. Et en toute logique, plus les informations étaient donc fiables pour les clients. Le point de départ est donc la donnée.

Un autre exemple est **Pandora**, une radio en streaming. Son offre est basée sur son « Music Génome Project » (MPG), un système breveté de catégorisation de chansons sur la base de 450 attributs et l'analyse de millions de titres permettant de définir des stratégies de radio calquées sur les goûts. Pandora permet de proposer des sélections musicales sur-mesure. Le Music Génome



Project **est sans possibilité** d'imitation pour une raison simple : les données sont totalement exclusives à l'entreprise. Les concurrents peuvent donc difficilement les répliquer. Ces données permettent le déploiement des systèmes intelligents comme le MGP. Et à ce titre, plus les données évoluent en elles-mêmes et plus la capacité à produire des connaissances grâce aux outils sera elle aussi meilleure.

→ ***Le quatrième point** est l'amélioration des utilisateurs en exploitant la qualité des données, puis la rapidité avec lesquelles les connaissances peuvent être intégrées aux produits, ce qui revient pratiquement au même. Spotify, Apple Music ont supplanté Pandora mais uniquement aux EU. Mais Spotify a connu un succès mondial car a progressé en exploitant une excellente base de données et en proposant ainsi des playlists avec un effet de réseau attractif, celui de pouvoir rechercher et écouter les stations de radio des autres utilisateurs. Ici des cycles d'apprentissage rapides font qu'il est difficile pour les concurrents de rattraper leur retard surtout sur les produits bénéficiant de plusieurs objectifs d'amélioration au cours du contrat type d'un client.

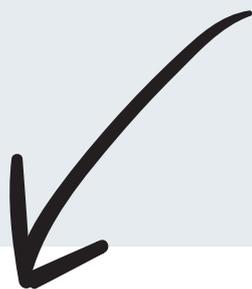


3. *L'intelligence artificielle pour les exploiter au mieux ex post*

L'intelligence artificielle, grâce au Machine Learning ou au Deep Learning, va permettre un second tour d'amélioration, le premier, comme démontré, concernait l'utilisation ex ante de la donnée, avant l'exploitation de celle-ci par un outil innovant. L'interprétation des résultats ex post lui confère un avantage indéniable dans le dessein de projets futurs, dans l'amélioration continue des exploitations ex ante des données et dans l'amélioration des premiers produits ou services délivrés. **Au-delà de ces points managériaux, il ne faut jamais oublier les aspects humains.** Cela vaut pour la technologie comme pour la data. L'art de la persuasion : face à l'essor de l'analyse des données, les entreprises ont engagé les meilleurs data scientists qu'elles pouvaient trouver mais la plupart du temps n'ont pas réussi à retirer la valeur exemptée de leurs initiatives en sciences de données. Pour qu'un projet crée de la valeur, l'équipe doit poser les bonnes questions, traiter les données pertinentes et en tirer les résultats, comprendre et communiquer ce que les résultats signifient pour l'entreprise. Il est très rare qu'une seule personne maîtrise ces deux domaines. La plupart des data scientists sont formés au premier mais pas au second.

Par Pascal de Lima,
Chef Économiste et Directeur

SAISIR L'AVANTAGE COMPÉTITIF : LES DONNÉES COMME LEVIER D'OPTIMISATION POUR LA GESTION DU RISQUE OPÉRATIONNEL



Les données sont, entre autres enjeux, un allié incontournable pour le pilotage du risque opérationnel, que ce soit en matière de prise de décision, monitoring, de respect des exigences prudentielles, de reporting. Pour tirer un avantage compétitif de ces données il faudra les valoriser au travers d'un modèle de données de risque opérationnel incluant digitalisation des process, urbanisation (infrastructure et architecture) et utilisation des techniques analytiques capables de livrer de l'information à haute valeur ajoutée aux acteurs clés. De surcroît, il faudra se doter d'une gouvernance et d'une capacité d'évolution en fonction des résultats de contrôles afin de pouvoir réagir efficacement et anticiper les événements majeurs et inconnus, sans oublier par ailleurs de répondre aux exigences du régulateur.

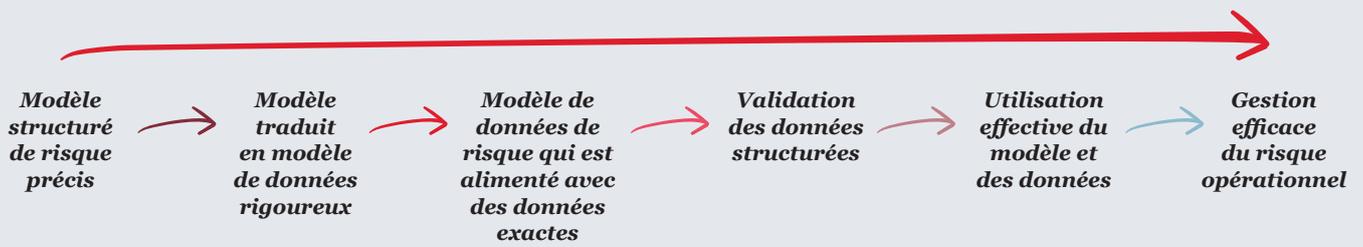
L'objectif vise à disposer de tableaux de bord complets, exhaustifs et dynamiques, prenant en compte l'appétit de risque de la banque apportant une visibilité sur le profil de risque présent et futur, tout en remplissant les exigences en matière de fonds propres, et de reporting.

Pour cela, il s'agit de mobiliser et de coordonner des Golden Source et référentiels nombreux, de créer une architecture et des flux de données qui permettent de cerner tous les risques connus et potentiels, puis de les analyser et de les suivre efficacement afin de connaître l'exposition en temps réel et ce, de manière dynamique.

LE BON MODÈLE DE DONNÉES

Pour prendre les bonnes décisions il faut quantifier correctement le risque, et pour cela, définir les mesures, process de collection, collecter les données et fournir des reportings et accès en temps réel.

Les principaux écueils à éviter à ce stade sont : un modèle incorrectement défini, l'utilisation de mauvaises données, ou encore un processus opérationnel de données défaillantes. De plus, un usage correct (adapté aux besoins et respectueux de la réglementation) doit être assuré et monitoré.



Un premier pas pour saisir l'avantage compétitif des stocks géants de données possédés par les banques est donc de développer un modèle de données adéquat et complet.

Un bon modèle reflète exhaustivement toutes les données requises, les sources, leur cheminement, les relations et dépendances entre les entités pour une gestion efficace du risque opérationnel.

Les océans de données de risque opérationnel englobent les cartographies de risques récurrents et exceptionnels, leur quantification, les processus d'affaires afférents, les plans de contrôle des trois lignes de défense et les résultats associés, y compris les exercices d'autoévaluation et les Key Risk Indicators. Viennent s'ajouter des données sur les recommandations, incidents et pertes historiques et externes, plans d'action, entre d'autres catégories. Ces données sont transversales en termes de fonction, d'activités et d'écosystèmes. Mapper cette complexité est logiquement incontournable.

Il faut donc analyser le contexte stratégique et les nombreux besoins pour déterminer et décliner les concepts, catégories et leur granularité. Par exemple, des techniques telles que l'analyse des facteurs de l'information sur les risques nous permettent d'analyser un risque en actif à un risque associé, événement de menace, parties prenantes primaires et secondaires, les scénarios et événements de perte, la fréquence et magnitude de perte, les contrôles de prévention et mitigation, etc.

Il faut également considérer les arbitrages entre la complexité du modèle et l'effort pour le maintenir et le valider. Le modèle doit ensuite être testé, par exemple, à l'aide des approches telles que les Arbres de défaillances (déductif, de la cause à l'effet), les Modes de défaillance et de leurs effets (Cf. encadré définitions), permettant d'assurer la qualité et de palier aux potentiels problèmes sur la data des risques. Le modèle devra être accompagné d'un cadre d'évaluation des capacités de données qui s'intéresse à la gouvernance, l'architecture, la modélisation, la qualité, les flux, le data warehousing et Business Intelligence, la gestion des référentiels et la metadata. Enfin, un usage correct de tout l'édifice requiert la mise en place de process, procédures et d'une gouvernance dédiée.

TECHNIQUES TRADITIONNELLES ET IRRUPTION DES TECHNIQUES ANALYTIQUES AVANCÉES

Une fois dotées de systèmes permettant l'accès à la donnée et sa maîtrise, les banques doivent ensuite se poser la question des techniques d'analyse appropriées. Selon le besoin, on pourra mener une identification de risques, des causes ultimes et de ses interactions, quantifier les risques, et de plus en plus, prédire un événement pour aider la prise de décision. Utiliser toute cette panoplie de nouvelles techniques est nécessaire pour se démarquer des concurrents.



Traditionnellement, la gestion du risque opérationnel dans le secteur financier est conçue comme un problème de classification binaire, ordinal (impact faible / conséquent), ou encore par catégorie. Pour l'élaboration des matrices d'aversion, le risque est qualifié selon les critères de fréquence et la cotation des impacts financiers et non financiers. On utilise des méthodes classifiant les risques dans des catégories non continues résultant en matrices de risque axées sur la probabilité-gravité/sévérité. Cependant, le risque opérationnel est mieux modélisé sur un continuum temporel et de sévérité. Par conséquent, une approche visant à les quantifier comme une probabilité continue fournirait une meilleure représentation des opérations et risques.

Par ailleurs, les banques utilisent des modèles analytiques dits "traditionnels", tirant des inférences sur la population à partir d'un échantillon basé sur la théorie des probabilités, souvent imposées par le régulateur. Citons à titre d'exemple les exigences de fonds propres à titre de risque opérationnel du cadre Bâle, comme les modèles Monte Carlo VAR ou l'approche de distribution des pertes LDA. Ces techniques ont un caractère de haut niveau descriptif et rétrospectif, elles reposent sur des données internes historiques sur les pertes, mais ne tiennent généralement pas compte des facteurs opérationnels connexes. Par exemple, les méthodes VAR fournissent une estimation d'un point particulier de la distribution. Elles deviennent aussi rapidement très complexes sous les contraintes empiriques, méthodologiques et réglementaires imposées, par exemple, concernant les sévérités extrêmes (heavy-tailed), c'est-à-dire les événements de perte à forte sévérité mais à faible probabilité, car les banques disposent rarement de données adéquates pour modéliser la queue de distribution. Un autre problème est la difficulté d'évaluer le niveau de corrélation entre les différents types de risques. En effet, ces modèles sont basés sur des données macro, dont le manque peut entraîner des difficultés importantes.

Dans l'approche de distribution des pertes - pour les institutions utilisant l'approche de mesure dite "avancée AMA" - des scénarios reflétant l'environnement commercial et les systèmes de contrôle interne sont évalués à l'aide d'opinions d'experts et à partir des données internes et externes sur les pertes. Cette démarche vise à obtenir une évaluation quantitative approximative des distributions de fréquence et de gravité des pertes opérationnelles sur un horizon donné. Le problème : bien qu'elle soit un outil puissant pour estimer les pertes, l'analyse de scénarios reste très subjective...

D'autre part, les banques mènent souvent des analyses sur les événements des risques ciblés, qui ne tiennent pas compte des interrelations complexes entre les facteurs techniques, sociaux et organisationnels. Or, si l'on souhaite obtenir un avantage en gestion de risque opérationnel, il faut élargir cette approche vers une approche "multirisque" qui considère en détail, plusieurs événements spécifiques. La crise du COVID illustre parfaitement ce que peut être une perturbation systémique impliquant une

La crise du COVID illustre parfaitement ce que peut être une perturbation systémique impliquant une contagion des risques.

contagion des risques, soulignant s'il le fallait l'importance des perspectives multirisques.

Néanmoins, de plus en plus l'industrie doit identifier des schémas prédictifs généralisables, c'est-à-dire réduire la dépendance vis-à-vis des experts afin de diminuer l'intervention humaine manuelle, limiter les biais humains et améliorer la cohérence.

Ainsi, pour appréhender et gérer efficacement le risque, il faut développer une vision stratégique intégrant le développement des compétences requises pour implémenter les techniques analytiques avancées. Selon les cas d'usage, on peut opter pour une approche d'analyse traditionnelle ou avancée, descriptive ou diagnostique (réactive et rétrospective), ou prédictive (proactive, aide à la prise de décision), ou encore prescriptive (décisions optimales).

L'ambition est de donner une orientation prédictive à la gestion du risque opérationnel grâce à l'apprentissage statistique et automatique moderne, IE., l'Intelligence Artificielle. La puissance des algorithmes permet d'extraire des informations plus approfondies avec moins de ressources.

Cela permet de mieux répondre aux besoins des clients ou de gestion de risque, d'acquérir davantage de compétences et de bénéficier des avantages d'économies d'échelle.

La bonne nouvelle est qu'il existe déjà une recherche avancée dans la quantification de l'incertitude et les dépendances de modélisation entre risques opérationnels au sein des organisations du secteur financier, notamment en matière de réseaux bayésiens. A contrario, des techniques statistiques traditionnelles, ces approches permettent de développer des systèmes d'apprentissage dynamiques qui détectent les fluctuations dans les profils de risque ou les facteurs de risque sous-jacents modifiant la probabilité d'un événement de perte. À un niveau micro, les algorithmes prédictifs ont présenté de bons résultats dans la première ligne de défense en appui du personnel « sur le terrain ». À un niveau macro, ils permettent d'envisager des scénarios inaccessibles jusqu'alors.

S'ADAPTER AUX NOUVELLES EXIGENCES RÉGLEMENTAIRES

Un levier supplémentaire est de renforcer stratégiquement les plateformes réglementaires existantes pour renforcer la gestion de risque opérationnel. Jusqu'à présent, les risques exceptionnels et à fort impact sont couverts par les fonds propres d'un point de vue prudentiel. Mais l'échéance réglementaire de janvier 2023 concernant l'implémentation de la nouvelle méthode SMA Bâle 4 en matière de risque opérationnel, entraîne une simplification du calcul des fonds propres au titre du pilier 1 (les pertes historiques seront neutralisées).

Toutefois, malgré cet assouplissement du calcul, les données de pertes devront être rapportées dans le cadre des templates COREP. De surcroît, en raison d'une exigence accrue en matière de transparence au titre du pilier 3, l'EBA a publié des guidelines concernant les préconisations à suivre par rapport aux bases de données de pertes. CCR3 inclut également des règles précises par rapport à la collecte de ces données (définitions de pertes brutes/nettes, seuils, exclusions).

Certains établissements ayant recours aux approches standard et avancée AMA possèdent d'ores et déjà des bases de données de pertes. Ces bases constituent une source riche contenant les dates, le montant, le type d'événement, les descriptions des causes et des conséquences. Ces informations sont stockées dans des sources structurées (typologies, contrôles, incidents) et non-structurées (rapports des pertes sous format libre).

Ainsi, pousser les exigences Bâle (huit lignes de métier multipliées par sept types de risques) pour optimiser leur exploitation au sein des modèles et systèmes de gestion de risque opérationnel, peut représenter un quick win pour les mettre à niveau, voire les renforcer, avec une transition agile de ressources réglementaires.



EXPERTISES MULTIDISCIPLINAIRES ET APPROCHE PROJET

Recueillir et gérer les risques opérationnels internes, externes et systémiques exige désormais de se doter de modèles et systèmes renforcés permettant de percer les complexes relations simultanées et interdépendantes.

Les techniques analytiques modernes et les approches multirisques élargissent la capacité d'expliquer, quantifier et prescrire une gestion optimale pour déceler quels sont les risques encourus, mais aussi pour les anticiper. De plus, les institutions ont souvent une base d'expertise et l'infrastructure pouvant créer des synergies pour développer de tels avantages compétitifs grâce aux données dont elles disposent d'ores et déjà.

Saisir l'avantage compétitif de la donnée et se préparer pour les nouvelles exigences réglementaires requiert des expertises multidisciplinaires et une approche projet qui peuvent faire défaut aux établissements. Que ce soit en matière d'élaboration des cartographies, de développement du modèle de données de risque opérationnel ou encore, du choix et déploiement de solutions analytiques, GRC ou réglementaires, Harwell Management est en mesure de déployer auprès de ses clients son expertise et ses équipes pour aider à atteindre les objectifs d'une gestion de risque optimisée.



Techniques et Définitions

Arbre de défaillances ou ADD

Technique déductive, de la cause à l'effet, qui met en évidence les relations de cause-effet, permettant de représenter les combinaisons possibles d'événements et de quantifier la probabilité d'occurrence d'un événement dit "redouté". Cette analyse permet d'identifier les scénarios conduisant à l'événement en amont et peut être utilisée comme un outil de diagnostic, prévoyant la ou les défaillances des composants les plus probables lors de la défaillance système.

Modes de défaillances et de leurs effets (FMEA)

Méthode inductive, de la cause à l'effet. Démarche systématique et proactive d'évaluation d'un processus pour identifier où et comment il pourrait échouer et pour évaluer l'impact relatif des différentes défaillances, afin d'identifier les parties du processus qui ont le plus besoin de renfort / changement.

Réseaux Bayésiens

Technique pour modéliser, mesurer et gérer le risque en utilisant une connaissance préalable des facteurs de risque causal et un raisonnement probabiliste, sous la forme d'un graphe constitué de nœuds et d'arcs orientés. Les nœuds représentent une variable qui a un impact sur le risque et dénotent des relations causales ou influentes entre les variables. Chaque variable du réseau se voit attribuer une distribution de probabilité sous-jacente basée sur des informations préalables. Les informations supplémentaires améliorent les estimations antérieures.



Techniques analytiques avancées de gestion de risque opérationnel

Objectif	Type	Techniques analytiques	Description de cas d'usage de gestion de risque opérationnel	Données en entrée
Identification des risques	Descriptif	Traitement automatique des langues (Natural Language Processing)	Exploration de texte pour identifier les risques pour construire des taxonomies et des référentiels de risque. Efficace et efficace pour identifier les risques actuels et émergents.	Articles de presse, états financiers, rapports de risque interne et d'audit.
	Descriptif	Regroupement de données (Clustering)	La détection d'anomalies Identifie les valeurs aberrantes et surveille les écarts par rapport aux conditions de fonctionnement normales, mettant ainsi en évidence les cas à risque les plus élevés (e.g. clients, transactions).	Dossiers des clients et des transactions, rapports de cas suspects de vol d'argent.
Facteurs de causalité du risque	Descriptif	Exploration de processus (Process Mining)	Identifie les non-conformités des systèmes d'information par le biais d'une analyse de processus et met en évidence les domaines d'amélioration des processus Analyses plus approfondies, moins gourmandes en ressources et moins invasives.	Flux de travail réels du SI, journaux des événements, processus métier, règles, audits et contrôles de conformité
	Diagnostique	Arbre de décision	Identifie les attributs et les comportements des employés qui menacent ou représentent des vulnérabilités de la sécurité des informations. Plus complet et plus efficace que les audits manuels existants de journaux échantillonnés au hasard.	Journaux de requête de base de données.
	Diagnostique	Réseau bayésien	Modélisation de plusieurs risques opérationnels et leurs facteurs de causalité à l'aide de réseaux bayésiens.	Données sur les conditions de l'entreprise, sur les événements de perte.
	Diagnostique	Exploration de données (data mining) : Règles d'association, Motifs séquentiels Réseau neuronal artificiel Cartes auto-organisées	Analyse des facteurs qui causent des événements à risques spécifiques dans des contextes spécifiques (micro-risque).	Données sur les événements de perte, rapports d'incident, données techniques du système (par exemple, inventaire et enregistrements de transaction).
Quantification des risques (1)	Prédictif	Ajustement de distribution paramétrique Théorie des valeurs extrêmes Échantillonnage empirique non paramétrique	Estimation du capital réglementaire avec diverses techniques, dont distribution des pertes LDA, pour paramétrer les distributions de fréquence et de gravité, ainsi que les structures de dépendance.	Données internes sur les pertes (y compris le montant de la perte, métier, type de risque et l'année), l'environnement des affaires et facteurs de contrôle interne.



Objectif	Type	Techniques analytiques	Description de cas d'usage de gestion de risque opérationnel	Données en entrée
Quantification des risques (2)	Diagnostique	Réseau bayésien	Estimation du capital macro/multirisque à l'aide de réseaux bayésiens. Cette technique incorpore intrinsèquement des dépendances entre les secteurs d'activité et les types de risques. Des effets d'avalanche sur les pertes attendues se produisent à partir de niveaux élevés de dépendances, permet de représenter les facteurs de risque idiosyncratiques pour chaque métier, les facteurs de risque communs à plusieurs métiers et la contagion des risques entre métiers. Estimations réalistes et reproductibles des pertes totales. Mais malgré l'approche systémique, uniquement certains événements sont incorporés, ainsi qu'un petit échantillon de facteurs de causalité.	Données sur les conditions de l'entreprise, sur les événements de perte. Jugement d'experts et données sur les événements de perte sur les procédures opérationnelles, les compétences du personnel, l'environnement de travail.
Prédiction des risques	Diagnostique Prédicatif	Règles d'association Détection d'anomalies	Prédiction grâce à une surveillance des conditions des facteurs de causalité identifiés à partir de l'analyse diagnostique. Une forme d'alerte précoce de détection des risques, souvent intégrée dans les systèmes de notification.	Données sur les événements de perte, données techniques du système.
	Prédicatif	Réseau bayésien Règles d'association Forêts d'arbres aléatoires Processus de hiérarchie analytique Réseau neuronal artificiel	Prévision et prédiction événements à risque à l'aide des facteurs de causalité.	Données sur les événements de perte, données sur le système technique, données environnementales, opinion d'experts.
Prise de décision sur les risques	Prescriptif	Méthodes de maximisation des récompenses Apprentissage par renforcement profond Théorie des jeux	Les systèmes automatisés de prise de décision cherchent à prescrire l'action / stratégie optimale, compte tenu de la connaissance des états passés, actuels et futurs prévus.	Données sur les sinistres, opinion d'experts.
Prise de décision sur les risques	Descriptif Diagnostique Prédicatif Prescriptif	Classification, prédiction Analyse des séries chronologiques Analyse des associations Détection d'anomalies	Cadres holistiques et qualitatifs pour gérer un risque opérationnel spécifique, avec une architecture de gestion des données, un module de construction de modèles aligné sur l'objectif de gestion des risques et l'interprétation et la visualisation des résultats. Développés pour un contexte spécifique et principalement testé sur un cas, limitant la généralisation. Nécessite d'une configuration manuelle considérable et de l'intervention d'experts.	Données sur les événements de perte, données du système technique, opinion d'experts.



Conférence AFCM

Le 29 Mars 2022, Harwell Management a eu l'opportunité de participer au séminaire annuel de l'AFCM – Arab Federation of Capital Markets. Patrick FOSSET, en qualité de Directeur de Mission est intervenu dans le cadre d'une table ronde autour du thème de la digitalisation du post-trade.

Cette intervention a permis de faire le point sur la situation actuelle mais également de revenir sur les éléments clés de la digitalisation : motivations, applications et facteurs de réussite. Au travers de différents thèmes abordés (finance durable, innovation, trading, etc.), Patrick Fosset a pu partager ses expériences et bonnes pratiques autour de use cases d'automatisation. Nous remercions chaleureusement l'AFCM pour cette opportunité et les différents intervenants pour ce partage d'idées et d'expériences, comme pour la qualité de ces échanges.

Pour visionner les interventions de Patrick Fosset, rendez-vous sur ce lien

<https://1drv.ms/vs!AnYFkieFvX2nhOZClvQ7Qi-WXiOf5A>

AFCM ANNUAL 20
CONFERENCE 22
29th - 30th of MARCH BAHRAIN



AFCM
ARAB FEDERATION OF
CAPITAL MARKETS

بورصة البحرين
BAHRAIN BOURSE



AFCM Annual Conference: Bahrain 2022

Conference Brochure

March 29th, 30th - 2022

Manama - Kingdom of Bahrain

Virtual

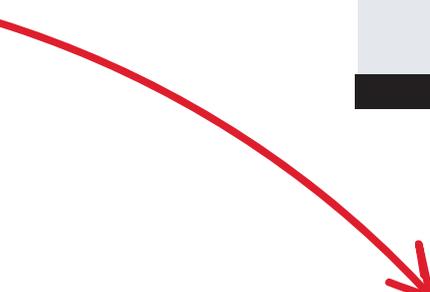
DirectFN

البنك الاماراتي المتحدة
ahli united bank

S/CO
سيكو

مباشر
MUBASHER

DIGITALISATION DU POST-TRADE : NOUVELLES TECHNOLOGIES ET DATA



L'automatisation des process de post-trade a toujours fait partie intégrante des activités liées aux Marchés de Capitaux. En quoi la situation est-elle différente aujourd'hui ?

Nous pouvons nous accorder pour dire que le principal changement visible est la rapidité avec laquelle les technologies évoluent, ainsi que la qualité et la capacité qu'elles offrent aujourd'hui en rupture avec l'amélioration continue pratiquée alors jusqu'ici.

Avant d'aller plus loin toutefois, revenons sur les principales motivations à l'origine de la majorité des initiatives d'automatisation des activités de post-trade sur les Marchés de Capitaux :

- * Améliorer la qualité des process en réduisant les erreurs manuelles (humaines),
- * Augmenter la capacité d'exécution des équipes de post-trade de façon exponentielle par rapport à la taille des équipes en question,
- * Renforcer la résilience des processus en environnement stressé,
- * Réduire les coûts / Améliorer le ROI.

Comprendre les motivations premières est déterminant afin de pouvoir différencier les projets d'automatisation à considérer comme des « must have », de ceux relevant plutôt du « nice to have ».

Une fois ces motivations internes identifiées et quantifiées, il convient de faire de même avec les motivations externes. Un besoin client clair a-t-il été exprimé ? Est-ce une requête d'un client, d'un régulateur, d'une contrepartie... ?

Dans tous les cas, la définition du use case doit rester une priorité, car il s'agit bien d'évaluer le coût d'opportunité d'un use case, et non celui d'une technologie. Trop d'initiatives d'automatisation ont malheureusement fini dans une impasse parce que les préoccupations premières étaient plus orientées "technologie", que bénéfiques réels et use cases.

Certaines initiatives ont néanmoins abouti : quelles sont les nouvelles technologies pour lesquelles des applications en post-trade ont été développées ces dernières années ?

LANGUAGE PROCESSING ET VISUAL RECOGNITION

Ces deux technologies sont le plus souvent appliquées au traitement de la documentation, par exemple pour la confirmation d'ordres. On retrouve notamment l'OCR (Optical Character Recognition) qui se démocratise.

BLOCKCHAIN (DISTRIBUTED LEDGER)

Encore peu déployée, elle fait l'objet de nombreux tests et POC par les différents acteurs, des Banques Centrales (autour des crypto-monnaies, CBDC – Central Bank Digital Currency) aux Métiers Titres (gestion du collatéral).

AUTOMATED MESSAGING WORKFLOWS

Principalement à destination des processus front-to-back impliquant plusieurs équipes/services, ils permettent de diffuser et partager l'information, ainsi que d'avoir des étapes de validation.

Il reste un dernier élément à prendre en considération, et non des moindres : la donnée.

Dans toute initiative d'automatisation, tout particulièrement si le process visé implique plusieurs intervenants sur le marché, la donnée est un composant central et inévitable.

Pouvoir échanger en utilisant un langage commun, c'est-à-dire une description partagée de la donnée,

est un prérequis pour automatiser un process, quel qu'il soit. Les travaux engagés par l'ISDA sur le Common Data Model pour harmoniser les données de marché sont un bon exemple de la criticité de ce sujet.

Au sein des organisations, l'angle majeur sera celui de la qualité de la donnée qui est primordiale pour espérer automatiser un process de façon suffisamment efficace pour en tirer des bénéfices. Garantir la qualité de la donnée suppose de passer par plusieurs étapes (dictionnarisation, contrôles, etc.).

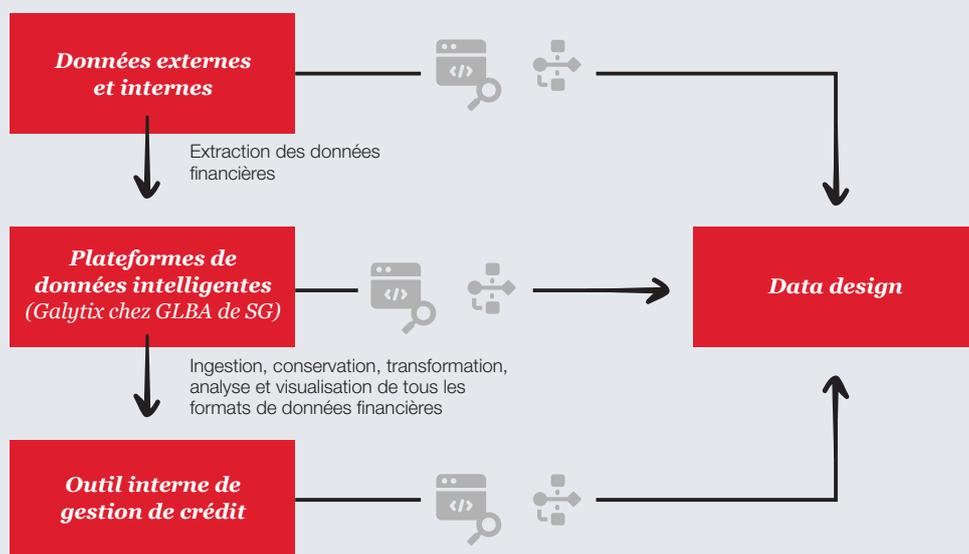
De nouvelles technologies, essentiellement basées sur l'IA, permettent également de mieux gérer et maîtriser la qualité de la donnée. Celle-ci peut alors devenir un actif de valeur pour l'organisation, et par extension, les processus liés peuvent aussi être générateurs d'efficacité opérationnelle.

Les enjeux de la digitalisation et autour de la donnée sont donc intimement liés.

La première a été accélérée avec la pandémie et la demande croissante de pouvoir tout faire à distance, jusqu'au sein des salles de marché. La seconde, si elle a toujours été présente, et a fortiori dans les activités de marché de capitaux, en devient par conséquent d'autant plus cruciale et prégnante qu'elle est sur le chemin critique de la première.

Par P. Fosset & A. Demangeon

Aperçu d'un projet OCR et DATA



I. Dans le cadre de l'octroi de crédit, les banques effectuent des analyses des données financières des clients. Pour cela, elles ont du mal à extraire des informations exploitables à partir de données internes et externes et doivent le faire le plus souvent manuellement.

II. Des solutions, comme l'OCR, permettent d'améliorer le process de collecte et d'analyse des données.

III. Des plateformes de données intelligentes permettent des décisions plus intelligentes, plus rapides et plus éclairées. La plateforme de Galytix intègre l'intelligence artificielle dans une architecture pilotée par pipeline - ingérant, conservant, transformant, analysant et visualisant tous les formats de données dans une seule solution.

IV. Afin de mener à bien ce type de projet, il est important de maîtriser les données concernées : le data design (dictionnaire et modèle) est donc essentiel, il permet d'avoir un langage commun.

OPPORTUNITÉS ET DÉFIS DES USE CASES DATA POUR LE SECTEUR FINANCIER

Dans un contexte économique très incertain, l'exploitation pertinente des données sur les clients, les produits et les marchés n'a jamais été aussi importante pour le secteur financier.

De nos jours, et grâce aux progrès des nouvelles technologies, il est désormais possible d'explorer, d'analyser et d'utiliser d'importants volumes de données qualifiées et ciblées. Cette exploitation à grande échelle des données permet de mettre en œuvre des cas d'usage métiers de plus en plus affinés.

Cependant, les investissements sont encore trop souvent réalisés sous un angle « **espérance de ROI** » plutôt qu'en « **ROI garanti** ».

Ainsi, quels sont les leviers à mettre en place au sein du monde des services financiers pour optimiser l'utilisation des cas d'usage data ? Les banques et les assurances sont-elles prêtes à franchir le pas ?

Optimisation de la relation client via l'ultra-personnalisation, segmentation de la clientèle plus fine et plus exacte, prédiction des besoins et optimisation des ventes, amélioration des calculs de probabilité de défaut, prédiction et lutte contre l'attrition des clients, fidélisation des collaborateurs, identification en temps réel des comportements atypiques pour lutter contre la fraude, aide à la décision sur les marchés financiers, audit en continu... le champ applicatif est colossal afin de permettre aux institutions financières de développer un avantage compétitif et différenciant dans ce nouvel écosystème de la data.



Certains cas d'usage sont devenus « classiques » mais la plupart relèvent encore de l'expérimentation sans garantie de résultat. Les raisons des échecs sur ces projets sont diverses : cas d'usage insuffisamment décrits, quantité ou qualité des données insuffisante, complexité de l'industrialisation, etc. Face à cet éventail de use cases possibles, les banques et les assurances n'ont tout simplement pas encore l'organisation et la culture adéquates pour tirer parti au maximum d'une bonne valorisation de leurs données.



Pour exploiter ses données de masse, une organisation financière dispose généralement d'un outil Data Lake, centralisant les données de différents systèmes sources, d'une solution de gestion des données, garantissant que tous les collaborateurs aient accès à des données fiables à tout moment, ainsi qu'éventuellement d'une solution de data virtualisation pour rendre accessible plus facilement les données sous forme de reportings prédéfinis.

Quand bien même le Data Lake constitue espace de stockage global et complet, il ne permet pas d'inscrire les données au sein d'un modèle métier et de servir un cas d'usage spécifique en phase avec les enjeux de sécurité, de qualité, de gouvernance et de traçabilité des données de l'organisation. Il en est de même pour un outil de gestion des données.

Quelles optimisations mettre en place ?

Tout d'abord, outre la nécessaire sensibilisation des collaborateurs à la data, l'adoption d'une philosophie essai/erreur avec le droit à l'échec (« try fast/fail fast »), l'acceptation d'arbitrages et de renoncements (toutes les idées ne pourront pas aller au bout) sont des prérequis indispensables au pilotage des cas d'usage data. Le métier doit être particulièrement conscient des risques inhérents à ce type de projets : en effet, pas de ROI « à coup sûr ». L'acceptation, culturellement et budgétairement, d'un certain taux d'échec s'avère indispensable et l'intérêt économique à arrêter le plus tôt possible les projets sans

avenir, se démontre de plus en plus. Il est possible de capitaliser d'un point de vue méthodologique sur les échecs et les réussites. En parallèle, un backlog alimenté en permanence d'idées de cas d'usage doit être maintenu en temps réel.

Une fois cette marche culturelle franchie, la première étape est sans doute la prise de recul nécessaire à la maturation de l'usage recherché. En plus de l'identification habituelle du besoin métier avec KPI business et estimation du ROI, s'ajoute une phase d'idéation data, assez nouvelle pour une institution financière. En effet, avant même de réfléchir aux premiers algorithmes, un cadrage du datalab s'impose : cycle « collecte de données, mise en forme et qualité, data mining, conformité et sécurité de la donnée » en méthode agile pour identifier au plus tôt les problèmes ou les impossibilités en amont (problème de qualité des données, récupération des consentements à l'utilisation de données à caractères personnelles, etc.).

C'est seulement une fois cette large phase de cadrage réalisée, que la recherche d'algorithmes proprement dite peut démarrer avec la réalisation du modèle par itérations jusqu'à performance souhaitée. Les KPIs algorithmiques adaptés en fonction des modèles choisis (matrice de confusion pour modèles de classification, MSE pour courbes de régression...) auront été préalablement définis.

Avant d'envisager un déploiement à grande échelle, un test du modèle en condition réelle d'utilisation avec retour d'expérience devra être réalisé via un ou plusieurs pilotes. Si cette mise en situation de production est validée, alors pourra s'ouvrir la phase d'industrialisation.



Cette phase d'industrialisation est très certainement la plus délicate pour les institutions financières. En effet, comment faire cohabiter au sein du legacy traditionnel des lignes de codes sources qui entraînent et réentraînent le modèle en temps réel ? Outre les recettes « IT » et « Métier » éprouvées sur des données de production, une méthodologie du développement du processus industrialisé de bout en bout doit être déployée (data preparation, feature engineering, écriture du code, refactoring production du score, exposition et IHM, sécurité du modèle et monitoring).

Enfin, la mise en marché du modèle s'accompagnera d'un packaging robuste autour de contrôles qualité, suivi de la performance, réentraînement régulier du modèle et gestion des évolutions du modèle.

L'intérêt sera donc de construire ce modèle de pilotage des cas d'usage data suivant une méthodologie éprouvée et structurée entre business owners, product owners et data scientists.

Au-delà de ces points managériaux et organisationnels, il ne faut jamais oublier les aspects humains. Cela vaut pour la technologie de manière plus générale et en particulier pour la data. Il est rare qu'un data scientist puisse à la fois poser les bonnes questions, puis traiter les données pertinentes, en tirer les enseignements et enfin comprendre et communiquer ce que cela signifie pour la banque ou l'assurance. Pour qu'un projet use case data crée de la valeur, il doit se structurer autour d'un trinôme business owner, product owner et data scientist et se traiter en mode projet agile, avec un delivery à 6 mois maximum.

Le contexte économique et réglementaire mondial, le besoin de personnaliser l'expérience-client et l'innovation véloce des Fintech imposent au secteur financier une transformation structurelle.

À ces défis s'ajoute la capacité à exploiter intelligemment un volume de données d'entreprise disparates mais de façon granulaire et sécurisée pour répondre aux besoins de l'ensemble des métiers.

Développer une méthodologie de pilotage des cas d'usage data innovants et agiles, tenant compte des réalités de l'organisation, et capable de suivre l'évolution des besoins métiers, permettra à chaque utilisateur de s'approprier rapidement son propre modèle métier pour à la fois observer et exécuter rapidement des données en se basant sur une approche itérative. Une agilité conçue autour du « besoin », indispensable au dynamisme et à la transformation du secteur financier.

Par I. De La Salle et P. De Lima

COMMENT TIRER UN AVANTAGE COMPARATIF DE LA DATA (APPLIQUÉE AUX STRESS TESTS CLIMATIQUES) ?

La réalisation de l'exercice pilote de 2020 a permis aux régulateurs européens de s'imposer comme fer de lance de la question environnementale et son impact sur les différents indicateurs financiers. L'exercice 2022 devra permettre de conforter cette position.

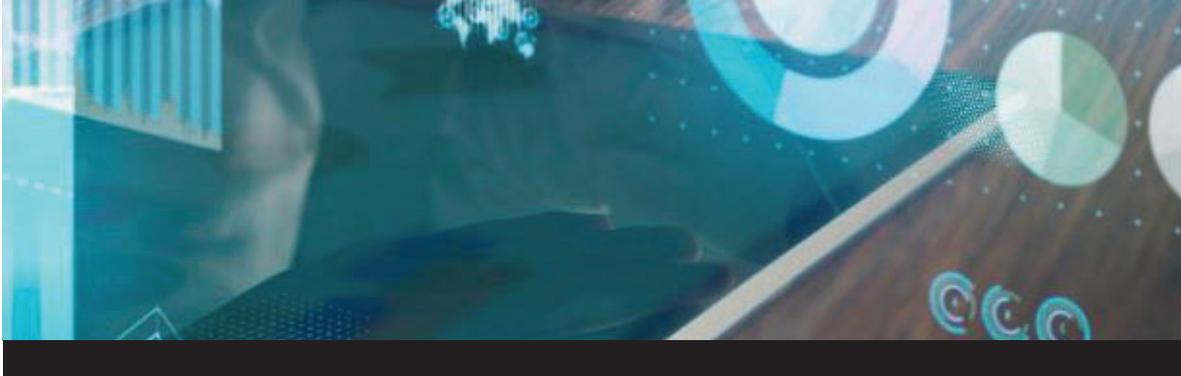
Néanmoins, les résultats de l'exercice de 2020 et publiés par l'ACPR soulèvent de nombreuses questions. En effet, plusieurs acteurs du secteur financier qui se sont engagés sur les enjeux climatiques considèrent que l'impact sur les différents risques semble faible, tout comme les niveaux d'exposition des portefeuilles de crédit. L'agence de notation Fitch Rating a également estimé que les résultats étaient débattables et que « les banques ont des trous dans leur rapport d'exposition et ont récolté très peu d'informations granulaires de la part de leurs clients ».



Globalement, les différentes exigences induites par les accords de Paris, étudiées au niveau bancaire par le prisme des exercices de stress, entraînaient la recherche de données toujours plus granulaires et souvent de niveau transactionnel. Chaque module regroupe un ensemble d'objectifs précis : émissions de GES, répartition des obligations au niveau des NUTS 3, etc... Chaque élément se base sur un socle de données que l'on qualifiera de « classiques » (date de début d'une transaction, encours, pays de résidence de la contrepartie...) et sera complété d'informations nouvelles : obligations classées, non pas face à leurs encours, mais face à leurs propres revenus, ventilation NACE, notations spécifiques... Ces dernières sont souvent non disponibles et demandent potentiellement la mise en place de proxys à l'image du rating CPE.

À ce titre, le développement des architectures SI, dans le but de fournir des informations toujours plus granulaires aux banques rapportées aux enjeux climatiques est une nécessité.

Les banques ont des trous dans leur rapport d'exposition et ont récolté très peu d'informations granulaires de la part de leurs clients.



DES EXIGENCES RÉGLEMENTAIRES À LA STRATÉGIE BANCAIRE

Le stress test climatique s'inscrit dans une logique d'exigences réglementaires toujours plus fortes. Là où les crises économiques passées ont entraîné un socle réglementaire toujours plus dense, celui-ci venait impacter la stratégie des banques.

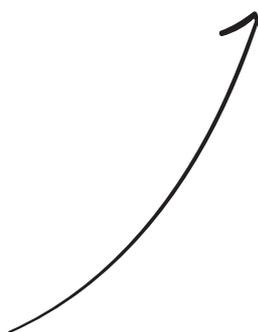
Par exemple : la crise financière de 2008 a accéléré les accords Bâle III, entraînant des exigences sur la Liquidité, avec l'introduction des ratios LCR et NSFR.

La nécessité d'une couverture de 100 % des sorties nettes de trésorerie par les actifs liquides a poussé certains établissements de crédit à revoir leur stratégie de gestion actif-passif, et sacrifier la rentabilité à la disponibilité des ressources. Cette stratégie a néanmoins un intérêt pour les banques : se conformer aux exigences réglementaires permet de prouver la solvabilité des banques, évitant un monitoring éventuel par les régulateurs, tout en renforçant leur crédibilité auprès des instances internationales, mais aussi de sa clientèle. Dans un climat d'incertitude, renforcer la confiance dans les institutions financières peut notamment permettre de renforcer leur « Goodwill » et par conséquent, potentiellement leur taille de bilan.

L'attention toujours plus importante apportée à la crise climatique ne fera pas exception et entraînera des exigences réglementaires toujours plus conséquentes. Le premier exercice obligatoire 2022 de la BCE servira certainement à l'élaboration future de contraintes réglementaires se rapportant aux lois Climat.

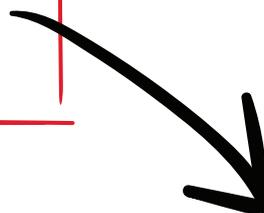
Ainsi, répondre aux interrogations levées par les résultats de l'exercice climatique pilote, en s'assurant d'avoir accès à des informations toujours plus granulaires sur les données climatiques pour remplir toujours plus efficacement les stress tests climatiques devra répondre à trois objectifs. Premièrement, rassurer les régulateurs en s'imposant comme établissement de référence sur les métriques écologiques. Deuxièmement, démontrer sa crédibilité en termes d'engagement sur les enjeux des accords de Paris, ce qui pourra indirectement permettre d'être créateur de valeur pour l'établissement, celui-ci faisant la différence en termes d'image. Troisièmement, réduire les coûts futurs liés à des productions d'informations réglementaires dans des délais contraints.

Répondre efficacement à ces exigences de contraintes réglementaires permet également de réduire les contraintes sur les collaborateurs bancaires. Certains développements s'avérant plus conséquents que d'autres. À titre d'exemple, la mise en œuvre d'Anacredit a mobilisé les équipes projet pendant plus d'une année, parallèlement à la mise en place très structurante de la norme IFRS 9. À ce titre, un Système d'Information plus performant permet de gagner en temps de développement et de réduire le coût en ETP.



Décryptage des données demandées dans le stress test ayant un enjeu data

Les 3 modules du stress test 2022 étudient les enjeux climatiques selon différents prismes, chacun possédant des enjeux au niveau des informations disponibles.



Module I : QUESTIONNAIRE QUANTITATIF

Le premier bloc est à caractère purement quantitatif et n'entraîne pas d'enjeux particuliers en termes de data disponible, les questions étant par ailleurs à menu déroulant, et ayant vocation à décrire certains aspects de la stratégie bancaire face aux enjeux climatiques. Pour autant, détenir davantage d'informations climatiques pourra affiner la stratégie bancaire face aux impacts écologiques.

Module II : MÉTRIQUES

Les deux métriques étudiées sont claires, l'objectif clairement défini : étudier l'apport des établissements financiers dans les émissions de gaz à effet de serre des entreprises. Néanmoins, les données nécessaires à la réalisation du module II peuvent être difficiles à trouver, les rapports annuels ne mettant pas l'accent sur les émissions de Co2. Cela peut conduire les parties prenantes du Stress Test à réaliser des proxys lorsque les données ne sont pas disponibles, expliquant à la BCE la méthodologie d'approximation (ce qui pourra conduire à des échanges lors de la phase de Quality Assurance).

Ce travail peut de plus nécessiter un grand nombre d'ETP, ce qui encouragera à envisager le développement futur d'informations relatives notamment aux émissions.

Les émissions GES

Pour identifier le bilan des émissions, les principes méthodologiques portent essentiellement sur deux points :

- * Le périmètre des différentes sources d'émissions
- * Les données à utiliser

Sur le périmètre, il est nécessaire de considérer le périmètre organisationnel (quelles filiales, quelles parties de l'entreprise sont considérées) mais aussi opérationnel (quelles sont les activités ciblées).

Sur les données, il est nécessaire de rassembler les informations déclarées par la structure étudiée. Il peut s'agir d'informations déjà disponibles dans l'organisation, mais aussi des données externes à recueillir auprès des fournisseurs, clients, usagers, etc... voire des données statistiques ou encore d'études (données génériques ou secondaires).

Une source externe existe et permettra d'avoir des informations nécessaires à l'estimation des émissions GES, à savoir « Base Carbone ». Les données permettent ainsi d'établir un bilan en tonnes de Co2 (ici, l'exemple pour l'entreprise Total). Néanmoins les informations disponibles sont imprécises. Les banques peuvent également solliciter leurs clients, mais l'exercice 2020 a illustré que cette collecte pouvait être difficile.

Module III : PROJECTION DES STRESS TESTS

Risque de transition

Le risque de transition, étudié par le biais des risques de crédit et de marché est la conséquence de l'évolution de l'attitude des différentes contreparties des groupes bancaires face aux évolutions en matière de politique et de méthodes de production (sur le plan organisationnel, stratégique ou technologique) pour répondre au besoin de réduction des émissions de gaz à effet de serre.

Risque de crédit

Les éléments étudiés sur le risque de crédit comportent plusieurs enjeux en termes de data disponibles.

Rating CPE

La notion de Certificat de Performance Energétique (CPE) a pour objectif de déterminer la quantité d'énergie consommée par un bâtiment pour une utilisation normale (température optimale, lumière...). Il se base sur plusieurs facteurs comme l'isolation, la chaufferie etc..., et consistera à délivrer une note à chaque logement. Cette note est donc propre à chaque logement et pourra varier d'une note de A (très économe) à G (très énergivore). De fait, le rating CPE va constituer la « carte d'identité énergétique » des bâtiments et des logements.

Les établissements bancaires doivent être capables d'affecter leurs expositions hypothécaires selon leur rating CPE. Les établissements doivent également proposer une méthodologie d'estimation du rating si une partie de l'information était indisponible.

Cette notion n'est pas forcément disponible lors de l'octroi de crédit et devra systématiquement être reprise par les agences bancaires, puis être remontée dans les systèmes d'information.

De fait, les établissements bancaires auront donc intérêt à développer des outils internes en vue de décrypter l'exposition de leurs financements aux industries émettrices de gaz à effet de serre.



Code NACE

NACE signifie : Nomenclature Européenne pour les Activités Economiques des entreprises. Il s'agit d'une nomenclature détaillée. Il existe en effet 615 codes à quatre chiffres que l'UE détermine ; les États membres attribuant un cinquième chiffre pour se distinguer par zone. Dans le cadre du Stress Test, les codes ont été limités à 22 macro-catégories.

Les établissements bancaires ne ventilent pas forcément leurs expositions selon cette nomenclature, d'autant que les codes NACE utilisés dans le Stress Tests constituent eux-mêmes un agrégat par rapport à l'ensemble de la nomenclature européenne. De fait cela demande un travail éventuel d'adaptation du Système d'Information.

Sur la définition des stratégies

Les stress tests climatiques ont vocation à déterminer la stratégie bancaire face aux enjeux écologiques. Ces stratégies sont basées sur une appréciation des membres de la gouvernance de chaque établissement, déclinée au sein des différentes directions. La part qu'y occupent les enjeux climatiques va donc dépendre partiellement de l'appétence de chacun aux enjeux environnementaux qui affectera le modèle stratégique retenu. De fait, rassembler de l'information sur la part de l'enjeu climatique dans la stratégie peut s'avérer ardu.

Néanmoins, la BCE s'attendra à retrouver des informations au sujet de :

- * l'impact en matière de stratégie pour chaque banque de l'enjeu climatique, et ceci de façon granulaire. Les établissements doivent prendre en compte :
- * la capacité de chaque client de changer son modèle (notamment pour les entreprises brunes), et de faire évoluer son mode de production en tenant compte des exigences sur les émissions
- * le maintien de la relation. En effet, les établissements pourraient rediriger leur financement vers des entreprises plus à même de respecter les engagements pour le Climat.
- * le risque en termes de reconnaissance globale, et notamment si l'établissement a réputation à s'engager sur des financements verts/bruns.
- * la volonté de s'imposer comme un acteur de référence face à certains groupes ou à l'inverse de s'émanciper de certaines relations, que ce soit par secteur ou zone géographique.

À ce titre, la construction d'une base de données aidant à la prise de décision face aux enjeux climatiques pourrait constituer un avantage pour les établissements.

Sur la longueur des horizons de projection

Le risque de transition est étudié par le biais de deux horizons, dont un long terme (30 ans), avec une hypothèse de bilan dynamique dès 2025. Comme évoqué précédemment, modéliser une stratégie bancaire globale s'avère délicat, à plus forte raison sur une durée aussi étendue avec autant de variables et d'incertitudes introduites. Réaliser le stress test peut donc s'avérer être un travail très dense. À ce titre, développer la data sur le stress test (et notamment sur la comparaison de résultats de projections passées) auprès de différents exercices au développement avéré permettrait de modéliser avec toujours davantage de précision l'impact d'une forte évolution stratégique sur le long terme.

Risque de marché

Le risque de marché ne comporte pas spécifiquement d'enjeux sur la donnée. Le calcul se faisant grâce à l'utilisation des modèles internes de la banque pour estimer l'impact du choc du prix du carbone sur la juste valeur des obligations et des actions ainsi que sur le portefeuille de couverture, en incluant les variables fournies par les scénarios.

Risque physique

Lors de l'exercice pilote effectué en 2020, des difficultés ont été rencontrées pour évaluer le risque physique. Tout d'abord, il a pu sembler ardu d'identifier les expositions concernées par les différents risques physiques, ceux-ci dépendant de leur situation géographique mais aussi de l'historique des risques concernés. Or la répartition géographique de l'ensemble de la production d'une entreprise est souvent beaucoup plus complexe que l'information détenue au niveau d'une contrepartie, l'étude des risques sur les indicateurs financiers classiques se faisant traditionnellement davantage

état par état. La donnée nécessaire à la complétude du stress test climatique peut donc s'avérer difficile à collecter sur la base des architectures de systèmes d'information existants, mais des travaux sur la donnée pourraient permettre de résoudre ces manques.

Risque de sécheresse

Sur le risque de sécheresse, la prise en compte d'un scénario canicule au premier janvier 2022 implique d'être en mesure d'en calculer directement l'impact sur les expositions.

Or, à moins d'appliquer des proxys en se basant sur des canicules passées (exemple en France : canicule de 2003), il apparaît difficile de mesurer l'impact sur les scénarios de production sans avoir recours à des informations extérieures. La construction d'une base de données portant sur l'impact passé de ce genre d'évènement pourrait permettre d'effectuer des estimations satisfaisantes dans des délais contraints.

Risque inondation

Sur le risque inondation il est demandé d'effectuer une ventilation selon le niveau de risque inondation, celui-ci dépendant de la zone géographique NUT S3. Deux soucis se posent ici.

Tout d'abord, les établissements ne posséderont certainement pas à leur niveau des informations sur le risque inondation par zone, demandant un recoupement par zone avec des sources d'information externes, ce qui pourra s'avérer difficile. Tout comme pour le risque sécheresse, d'une base de données portant sur l'impact passé de ce genre d'évènement pourrait permettre d'effectuer des estimations satisfaisantes dans des délais contraints.

Par ailleurs, les établissements disposent souvent dans leurs bases de données, d'informations sur les contreparties en termes de code nature, client, et localisation, mais celles-ci restent souvent celles du siège de l'entreprise. Le découpage NUT S3 est extrêmement granulaire, ce qui implique soit de faire des estimations, soit de construire une base de données reposant sur les adresses des sites de production.



Conclusion

En résumé, les différents modules du stress test climatique font appel à diverses informations qui ne constituent pas des données remontées de façon classique sur les expositions, et sont appelées à être toujours plus granulaires. L'adaptation des systèmes d'information pour fournir des données climatiques toujours plus granulaires permettra aux établissements, au-delà de remplir les reportings, de se doter d'un avantage comparatif en s'imposant comme acteur de référence du changement climatique aussi bien auprès des clients que des régulateurs, et anticipera la mise en place d'exigences futures.

Par C. Almenar et D. Cioffi

CONCLUSION DE CETTE DOUZIÈME ÉDITION



Pour conclure, il est certain que la data peut s'intégrer au cœur d'une multiplicité d'usages. Qu'il s'agisse de générer de nouveaux revenus grâce à la proposition d'expériences, de produits ou de services personnalisés, de mieux piloter ses risques, d'améliorer l'efficacité de ses processus métiers, de reporter de manière plus sécurisée et efficace aux instances de contrôle, la bonne utilisation de la data devient aussi évidente que le recours aux outils IT eux-mêmes lors de la phase d'expansion de l'informatique dans les années 80. D'une certaine manière nous redécouvrons les capacités offertes par le génie logiciel couplé à une utilisation/exploitation de données présentes massivement, fort de capacités de calcul, de stockage et de professionnels compétents démultipliés.

La phase actuelle de généralisation de l'exploitation intelligente de la data est loin d'être achevée, tant la data est intimement liée à toute production d'une fonctionnalité, d'un raisonnement et in fine d'un service ou d'un produit. La mise en qualité de cette dernière représente une difficulté, et donc un avantage comparatif. La créativité des équipes (métiers, data scientists, analystes, etc.) dans ce domaine comme dans d'autres restera également un avantage comparatif.

Tout cela ne doit pas éluder la prudente et raisonnable notion/question du ROI.

Certes nous pouvons extraire de plus en plus de connaissances de la data, mais cela reste un outil, et un outil en tant que tel ne suffit pas, encore faut-il apprendre à l'utiliser au mieux. L'organisation métier avec ses équipes, ses pratiques doit elle aussi évoluer. Afin, d'une part, de s'approprier les nouvelles possibilités/applications offertes par cet outil, et d'autre part, ne pas se laisser déborder par ces nouvelles possibilités en restant focalisée sur la valeur ajoutée offerte en définitive, par tel produit ou tel service aux clients finaux.

Des équipes métiers et data travaillant ensemble, en méthode agile, avec des cycles courts d'itération, priorisant l'utilité, le bon timing, la valeur ajoutée, les revenus générés et coûts associés, du produit/service que peut servir l'exploitation de la data compteront parmi les critères de réussite de tout projet autour de la data. Rien de nouveau donc, si l'on se réfère aux facteurs clés de succès de tout projet articulé autour d'une technologie.

Enfin nous sommes passés à un niveau de maturité supérieur dans l'exploitation de l'information, il s'agit moins d'une rupture et plus d'une évolution « naturelle » et attendue, et comme souvent, pour ne pas dire toujours, la technologie est prête avant les hommes et les organisations.



Chez CGI Business Consulting, cabinet de conseil majeur en France, nous sommes audacieux par nature. Grâce à son intimité sectorielle et à sa capacité à mobiliser des expertises diverses, CGI Business Consulting apporte aux entreprises et aux organisations des solutions de conseil audacieuses et sur mesure, pour une réussite stratégique et opérationnelle de leurs projets de transformation. Nos 1 000 consultants accompagnent nos clients dans la conduite et la mise en œuvre de leurs projets de transformation, dans une relation franche et de confiance, pour leur permettre de prendre les bonnes décisions. Fondée en 1976, CGI figure parmi les plus importantes entreprises de services-conseils en technologie de l'information (TI) et en management au monde. Elle aide ses clients à atteindre leurs objectifs, notamment à devenir des organisations numériques axées sur le client.



cgi.fr/conseil

L'audace par nature