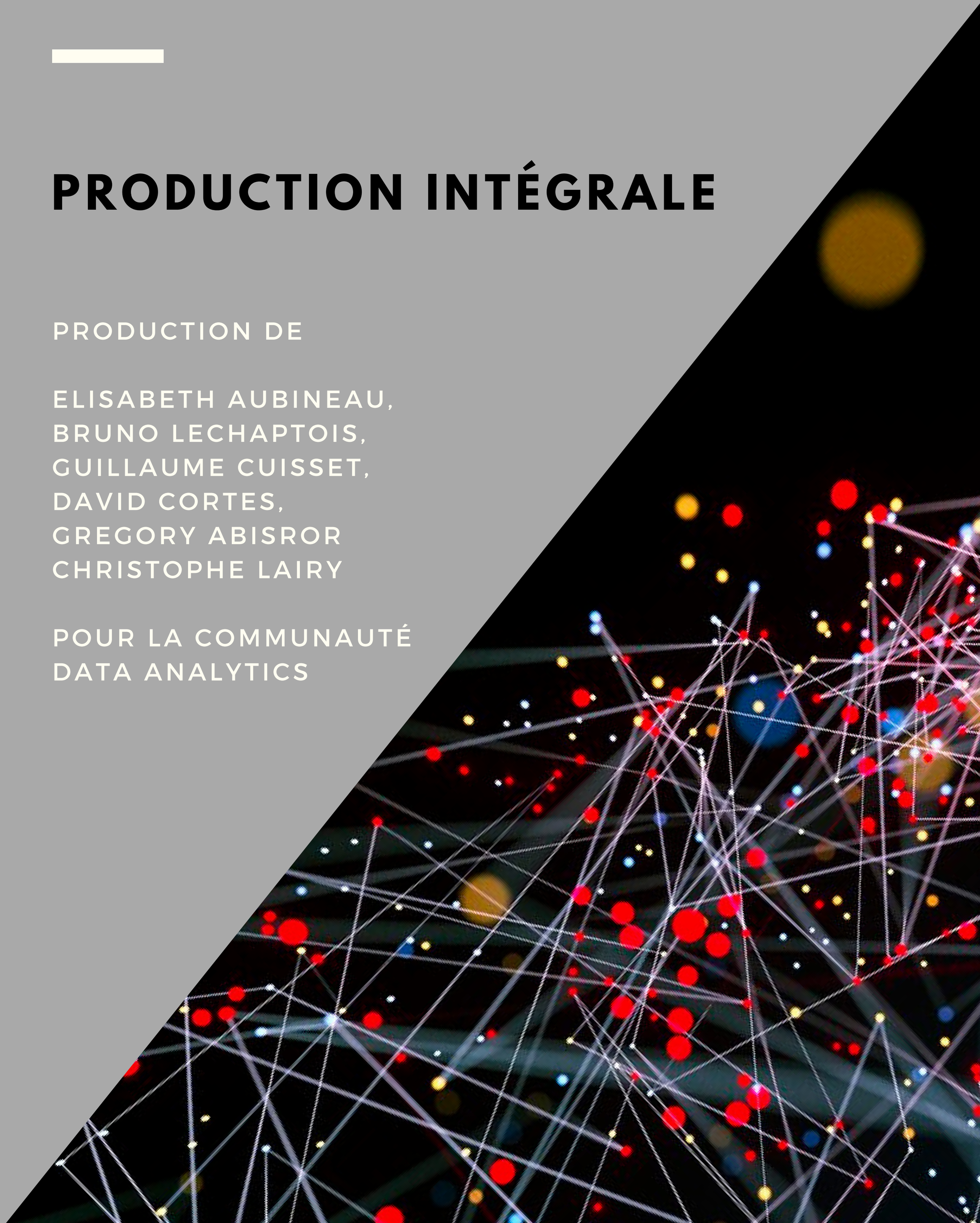
COMMUNAUTÉ DATA ANALYTICS

PRODUCTION INTÉGRALE

PRODUCTION DE

ELISABETH AUBINEAU,
BRUNO LECHAPTOIS,
GUILLAUME CUISSET,
DAVID CORTES,
GREGORY ABISROR
CHRISTOPHE LAIRY

POUR LA COMMUNAUTÉ
DATA ANALYTICS



1.

DÉFINITIONS DES DATA ANALYTICS

L’émergence du Big Data est considérée par certains comme une troisième révolution industrielle. Source de profonds bouleversements pour les organisations, de nombreux analystes s’accordent à dire que l’usage de la Data peut permettre de réduire les risques, de faciliter les prises de décisions et de rester compétitif grâce à la valorisation des données.

Voici ci-dessous des ressources qui permettent de définir cette notion plus en détail :

DÉFINITION DE DATA ANALYTICS

D’un point de vue très générique, le Data Analytics (ou Analyse de Données) est une science consistant à examiner des données, dans le but de tirer des conclusions à partir de ces informations.

L’analyse de données est utilisée dans de nombreuses industries et métiers depuis quelques années afin de permettre aux entreprises et aux organisations de mieux comprendre les flux passés, prendre de meilleures décisions et de plus en plus pour prévoir des événements futurs via l’analyse de données prédictive.

Deux dimensions sont généralement prises en considération concernant l’analyse de données :

- la donnée elle-même (sa nature, son volume, sa présence en entreprise et dans nos vies). La notion de donnée est considérée ici dans un sens large. Elle ne se restreint pas par exemple aux données d’identification ou de caractérisation des employés ou clients, mais est complétée par des sources externes (réseaux sociaux, bases de données publiques ou privées) aux formats très variés (vidéo, texte, « tweets », photos).

- Les traitements que l'on en fait (par exemple les algorithmes qui sont utilisés, leur complexité, leur interprétabilité).

Les sujets d’analyse de données se sont considérablement imposés dans les débats de la vie économique et de la vie publique. Deux tendances se sont renforcées dans la dernière décennie :

- Les données sont devenues omniprésentes, et archi-volumineuses (BigData ou grand volume de données).

- Les algorithmes sont devenus de plus en plus complexes (par exemple ceux basés sur l’analyse en réseaux de neurones) et sont développés dans une logique « apprenante » autonome, laissant à la machine une prérogative auparavant humaine.

Dans le prolongement de ces tendances, on constate actuellement qu’il y a deux phases disruptives et durablement installées :

- L'exploitation systématique des données, à des fins commerciales ou de contrôle.

- L'automatisation de certaines tâches humaines.

Des modes de traitement et d’utilisation de ces données se mettent en place dans des dimensions différentes, mais dans une logique de continuum autour d’une utilisation plus ou moins poussée de la donnée. Par exemple, pour simplifier il est possible de :

- Décrire les données récoltées, les classer et les analyser. Cette description peut être plus ou moins complexe selon le nombre de données traitées, et la complexité des traitements qui vont opérés pour obtenir la description voulue. Ainsi, la sémantique de Big Data est liée à ces processus de description mais pour un grand volume de données ou selon la complexité des traitements qui vont être nécessaires.

- Utiliser des algorithmes simples que tout un chacun peut maîtriser (ex. "arbre de décision", dont on fixe les valeurs sur la base d’un processus maîtrisé, et qu’on cherche à répéter ou à automatiser). C'est actuellement dans ce contexte que l'on déploie les robots. C’est de la robotique.

- Utiliser des algorithmes simples, mais moins contrôlables (ex. Random Forest, qui sont des arbres de décision, mais moyennés, et avec une dose de hasard dans l'apprentissage). C'est du "Machine Learning".

Enfin, dans le prolongement de la complexité, l’Intelligence Artificielle utilise tous les axes, et notamment la notion d’apprentissage.

2.

QUELS IMPACTS SUR LES MÉTIERS DE L'AUDIT ET DU CONTRÔLE INTERNE ?

QUEL IMPACT DU DATA ANALYTICS SUR LES MÉTIERS DE L'AUDIT ET DU CONTRÔLE INTERNES ?

Trois enjeux peuvent être considérés s'agissant du Data Analytics : l'impact pour les métiers du contrôle, les impacts que cette évolution apporte sur les modèles économiques et les processus des entreprises dans lesquels les fonctions de contrôle opèrent.

Pour le premier enjeu, les professionnels de l'audit et du contrôle interne, l'analyse de grands volumes de données se traduit par des opportunités de modifier les modes de travail en traitant par exemple des tests d'efficacité des contrôles de façon systématique et non plus par échantillon.

Pour le contexte de leur métier, cela implique aussi que ces fonctions de contrôle soient au fait des évolutions que l'analyse de données peut avoir sur l'entreprise et le contexte économique dans lequel elles opèrent.

En quelque sorte l'analyse de grands volumes de données est le premier maillon de la digitalisation pour l'entreprise et les fonctions d'audit et de contrôle.

Les attentes sont différentes et variées entre les besoins de l'opérationnel qui commence à s'investir dans la démarche pour ses besoins et les besoins des fonctions de maîtrise des risques qui doivent d'une part acquérir la démarche mais aussi avoir l'assurance que celle développée par les opérationnels est efficace et répond bien au besoin du dispositif de contrôle interne de l'entreprise.

La qualité des données utilisées et restituées et la revue de la démarche d'analyse développée sont ainsi des éléments des processus à réaliser pour permettre à l'auditeur ou au contrôleur interne d'utiliser les travaux des opérationnels dans leur évaluation des dispositifs de contrôle interne.

Dans une démarche basée sur l'utilisation de grands volumes de données, avec les outils technologiques puissants qui en sont le corollaire, on passe d'une analyse à base d'échantillons à l'intégralité des données et on regarde ce qui s'écarte d'une norme (jour, heure ...) selon des règles déjà connues. Les outils peuvent aussi effectuer des analyses statistiques pour constituer une référence **[1]**, identifier des transactions qui s'écartent des cas normaux, possibles erreurs ou schémas frauduleux. On passe alors de règles prescrites à de la recherche automatique, dans l'objectif d'une meilleure identification des zones de risques de l'entreprise.

L'analyse de grands volumes de données va se limiter à des démarches machines qui vont rester simples sans couche logiciel supplémentaire telles Business Intelligence, ou intelligence artificielle, Blockchain ou Robotics. La première problématique concrète sur lesquels des progrès ont été faits pour aider les processus des fonctions de contrôle est souvent le risque de fraude. Globalement c'est une première étape technologique dans la transformation des fonctions audit et contrôle.

Ainsi par exemple par rapport aux règles métiers, le seuil peut être mis par la machine. Le système peut également proposer des règles qui se rapprochent des règles frauduleuses identifiées et on cherche alors des analogies sur les types d'écritures par exemple, ou encore, le système identifie des corrélations entre variables et la machine va trouver des similitudes mais celles-ci ne sont pas forcément faciles à interpréter.

Pour aller plus loin :

- " Franchir un cap dans l'utilisation des données" - Étude PWC
- " Audit Analytics and Continuous Audit" - 2015
- " GTAG - Comprendre et auditer le Big Data
- " Data analytics is a game changer, but internal audit groups are lagging according to new study by Protiviti
- " La Data Analytics au service des entreprises et des métiers du chiffre

3.

DES ILLUSTRATIONS DE MISES EN OEUVRE PRATIQUES DES DATA ANALYTICS

Nous l'avons évoqué dans les chapitres précédents, les définitions de ce que recouvre l'analyse de données est vaste et variable selon les périmètres de leur utilisation. Cette variété de définitions implique une multitude d'applications et d'usages et ce sujet évoluera nécessairement avec la maturité des fonctions de contrôle et le déploiement d'outils idoines. L'acception la plus répandue est que du point de vue des « business models » des entreprises, les analyses de données en grand volume vont conduire par exemple à une meilleure connaissance des clients sur leur mode d'usage des services et produits qu'ils consomment. De là, un meilleur ciblage de l'action commerciale pour développer les ventes. On voit bien dans cette illustration les enjeux métiers et pourquoi certains auteurs considèrent que la data est l'or du 21eme siècle.

Parallèlement les contraintes légales et réglementaires s'accroissent pour tenter d'anticiper les débordements liés à l'utilisation de la donnée et protéger le consommateur lambda. Ainsi les règlementations sectorielles (Solvabilité ou Bâle), celles qui ciblent la lutte contre la corruption (Sapin 2) ou la cybercriminalité, et enfin celles centrées sur la protection des données personnelles (RGPD). Ces contraintes influent fortement sur les moyens déployés sur la gestion des données dans l'entreprise.

Les métiers du contrôle et de l'audit ne sont pas épargnés par cette tendance. En dehors du suivi des enjeux réglementaires et opérationnels, ils peuvent tirer des bénéfices des techniques d'analyse de données de grands volumes. Ainsi l'exploitation maitrisée d'un ensemble important de données peut permettre de remplacer des évaluations par sondage, beaucoup plus pauvres en information, par des évaluations s'appuyant sur une base de données complète. Le traitement des informations collectées peut ainsi renforcer l'efficacité du dispositif de contrôle interne dans son ensemble, et par là-même l'efficacité des fonctions qui sont en charge de superviser le dispositif.

Ces évolutions ne sont pas sans limites. La collecte et l'accès à de grands volumes de donnée implique plus de transparence, d'enjeux de qualité et d'exhaustivité, mais aussi une nécessaire gestion de la confidentialité et de la sécurité des informations collectées et traitées.

Par-delà les aspects techniques, c'est également une modification de la culture de la donnée au sein des entreprises (comme de la société) qui devra se mettre en place, avec des impacts aussi sur les organisations, qui pourront se structurer autour de ce traitement des données.

COMMENT LES FONCTIONS D'AUDIT ET DE CONTRÔLE PEUVENT-ELLES INTÉGRER CE NOUVEL ESSOR DE LA VALEUR AJOUTÉE DE L'ENTREPRISE ?

Une première réponse réside sans doute dans le fait de pouvoir donner une « assurance raisonnable » et raisonnée de la fiabilité des informations que les métiers ont à disposition pour répondre à leurs besoins opérationnels et aux aspects réglementaires ou légaux. La problématique des fonctions de contrôle est dans ce cadre la complexification des processus liés à la donnée et l'augmentation du nombre d'informations.

Comment apporter l'assurance d'une information fiable ?

Une information fiable signifie notamment qu'elle est vérifiable, reproductible, comparable et son accès contrôlé. Une information est fiable par rapport à un objectif qui, pour les fonctions de contrôle, est la qualité d'un dispositif de contrôle interne. On peut ainsi, par exemple, recommander pour de telles données de définir un principe de qualité contrôlable et vérifiable basé sur :

- une documentation décrivant clairement notamment le risque que l'on veut couvrir, le périmètre sur lequel s'applique l'analyse, le traitement ou calcul, et le résultat.
- un test effectué avec un plan de test validé.
- un suivi en maintenance (sauvegarde, suivi des évolutions avec vérification de bon fonctionnement, contrôle d'accès au paramétrage, ...)

Comment apporter l'assurance d'une sécurité des accès aux données ?

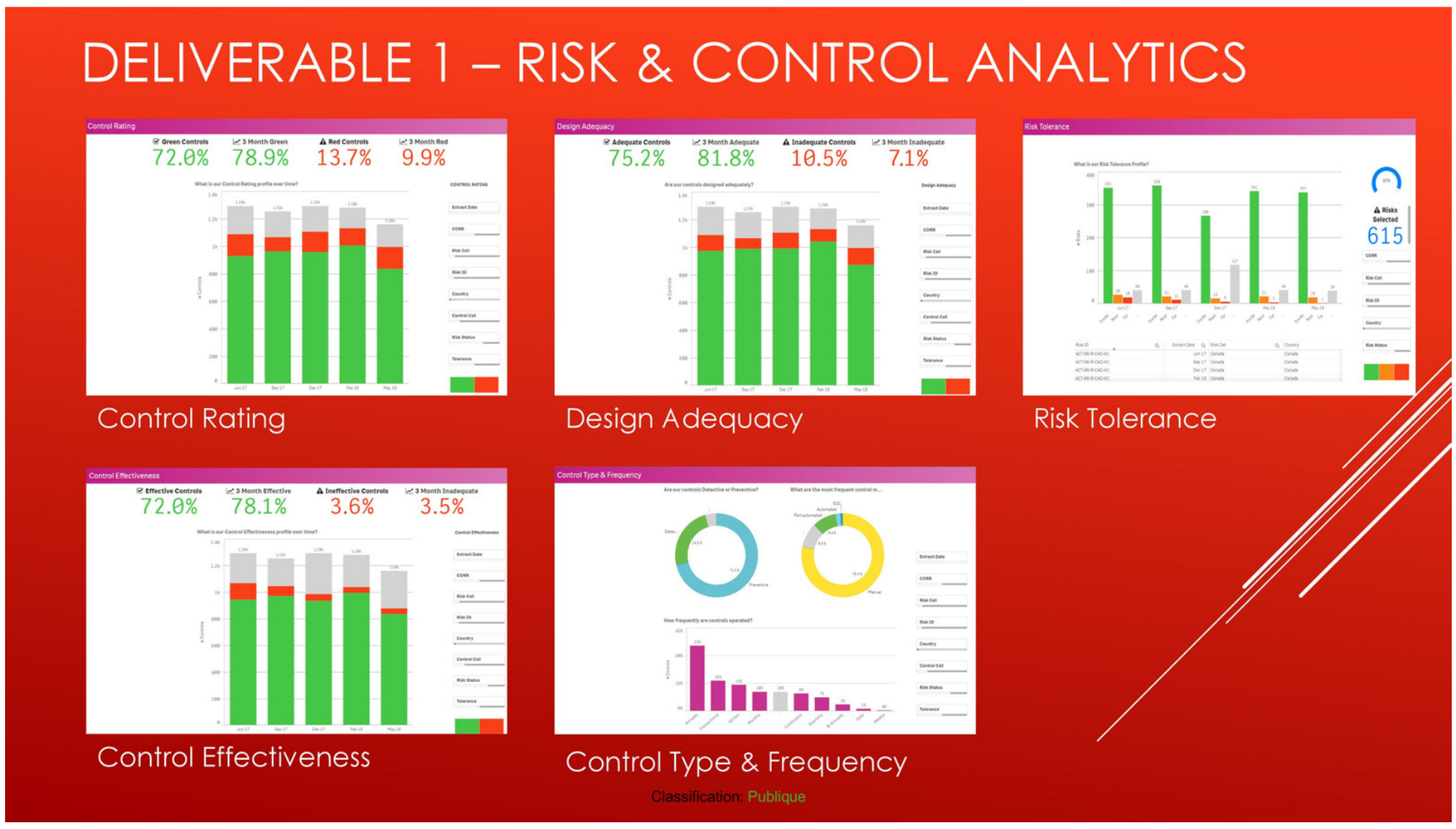
En dehors de l'aspect gouvernance des systèmes d'information qu'une culture des données implique, une assurance quant à la sécurité d'accès aux données sensibles passe nécessairement par de la technique. Des tests d'accès logiques ou physiques et de vulnérabilité sont indispensables à avoir dans le radar des fonctions de contrôle des risques et de contrôle interne. La fréquence de conduite de ces tests a de plus tendance à augmenter au fur et à mesure des innovations en matière de cyber-attaques. Il y a donc un impact non négligeable sur les ressources et le temps dédié à ces problématiques de sécurité des systèmes d'information et d'accès aux données sensibles.

Illustrations pour l'Audit Interne

Nous illustrons les bénéfices que l'analyse des données peut apporter pour l'audit interne sur les trois dimensions suivantes : analyse des risques, suivi des recommandations, analyse des recommandations.
L'analyse des risques et des contrôles doit permettre une optimisation de la couverture des risques entre les " lignes de défense ". C'est-à-dire structurer le " qui fait quoi " en première, deuxième et troisième ligne pour assurer qualité (des données financières en particulier pour l'audit interne) et sécurité des données.

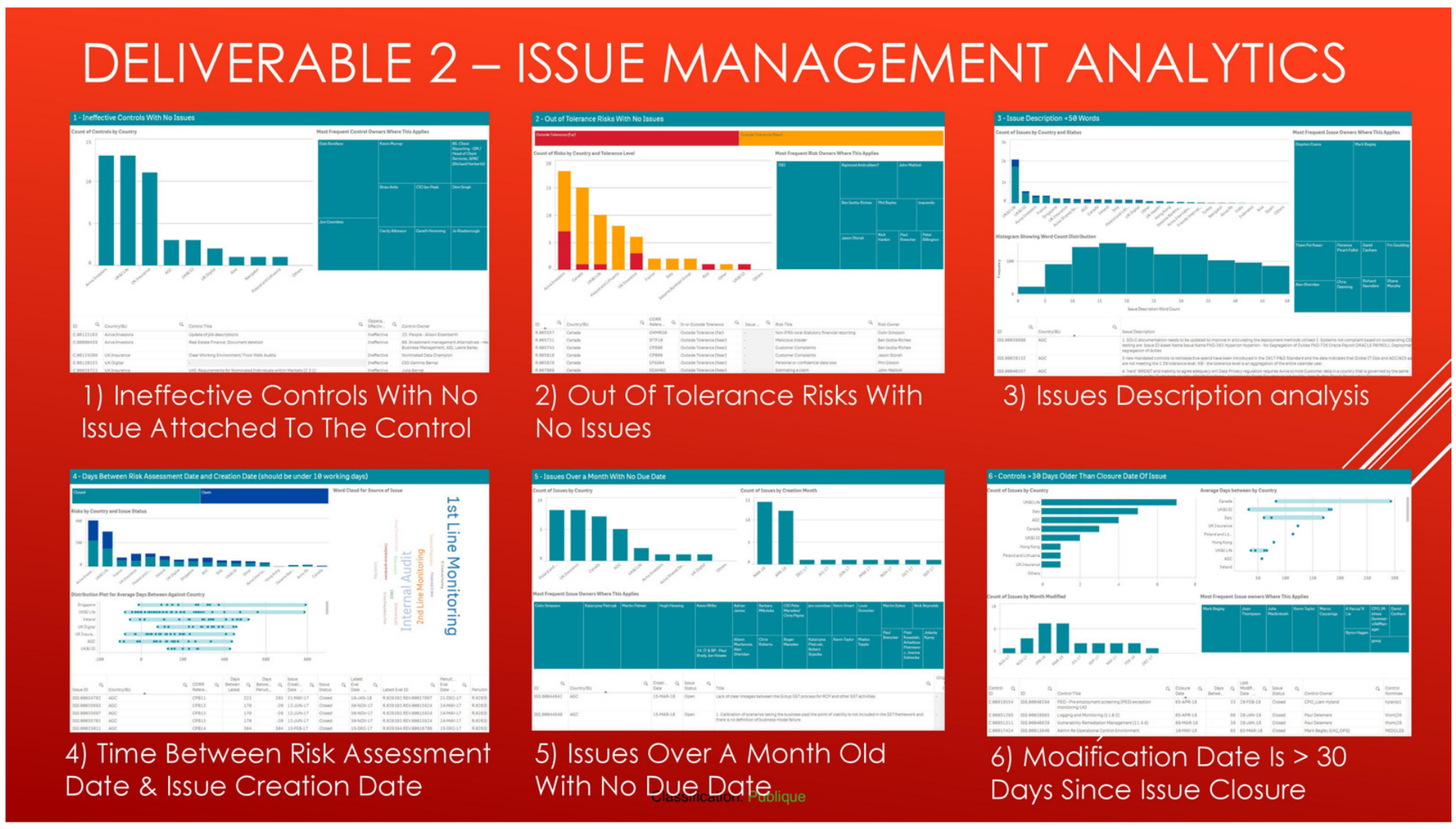
Analyse des risques et des contrôles

L'analyse peut se positionner sur les natures de risques, leur impact, leur risque d'occurrence. L'analyse peut se porter aussi sur l'efficacité des contrôles ou leur design. Mais cela nécessite déjà d'avoir un outil centralisé au sein de la société. Les graphiques ci-dessous décrivent pour l'évolution dans le temps de l'évaluation des risques en impacts (selon une échelle Vert-Orange-Rouge pour faible-moyen-important) ; l'évaluation de la qualité de design des contrôles ; l'évolution des tolérances des risques tels que fixés par le management ; l'évolution de l'efficacité des contrôles ou le type de contrôle (préventif, détectif, manuel, automatique) et leur fréquence (mensuel, quotidien, ...).



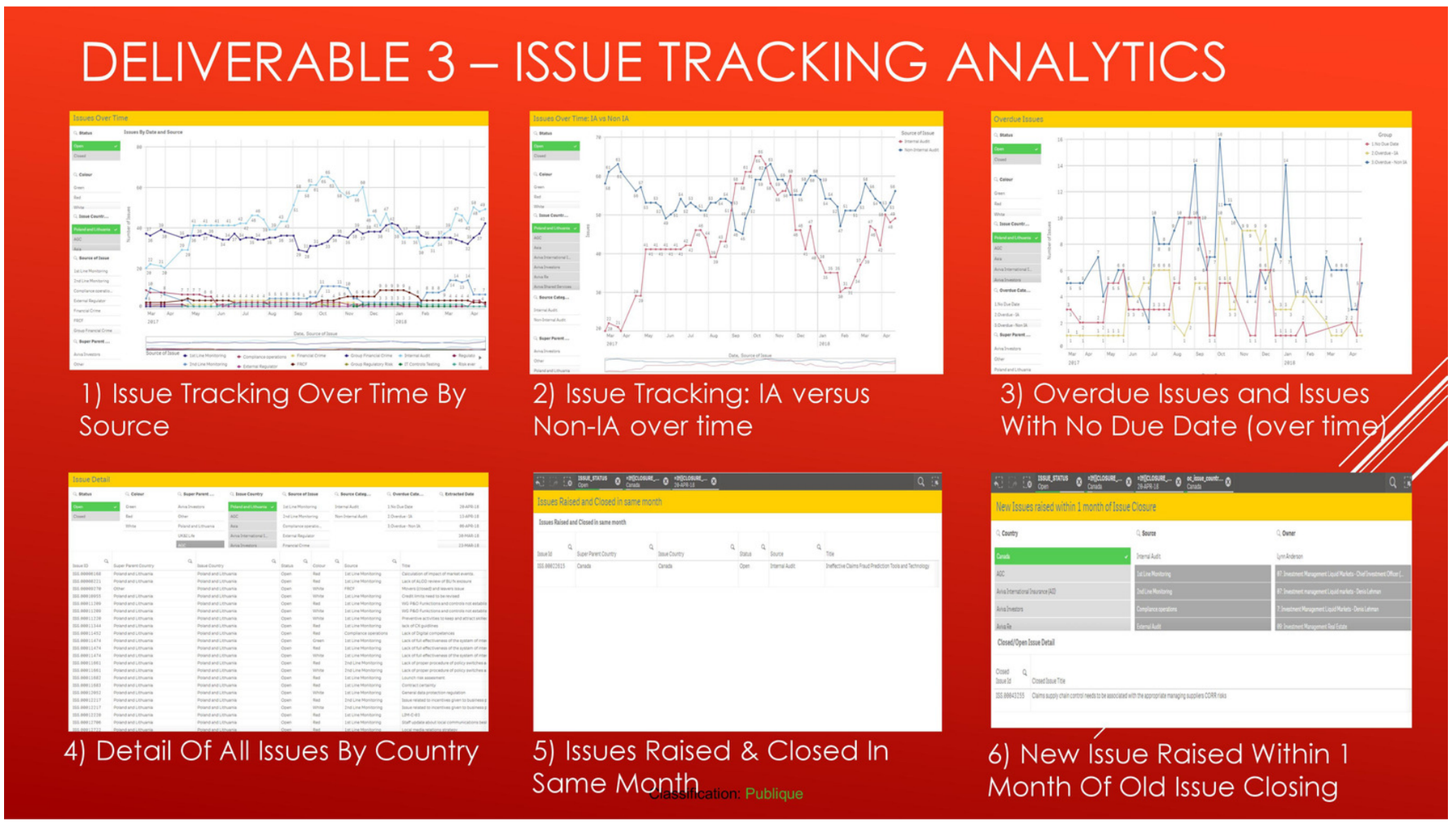
Analyse de suivi des recommandations

L'analyse des recommandations (fermeture dans les délais, répartition par département, délais de clôture, niveau de risques, etc..) est un processus souvent consommateur en temps pour les équipes d'audit interne.
L'analyse de données peut optimiser ces processus et permettre aux auditeurs de se centrer sur l'analyse des documents communiqués pour la clôture des recommandations plutôt que sur les reporting de suivi ou les échanges de documents. Là encore, l'existence d'un outil partagé avec les autres départements sur la collecte des évidences de contrôle est essentielle à la qualité des analyses qui seront faites. Nous donnons ci-dessous des illustrations de suivis. Les possibilités sont nombreuses pour le suivi des recommandations et pour s'assurer que les opérateurs comme les managers et les fonctions de contrôles mettent en place dans les délais et de façon appropriée les recommandations. Les exemples ci-dessous ne sont pas exhaustifs et chaque société doit avoir ses propres règles pour assurer un reporting approprié vers les organes de gouvernance.



Analyse des recommandations

En complément des suivis réalisés pour des aspects reporting ou de suivi des fermetures de recommandations. Une analyse des données peut permettre une vision avec les risques de l'entreprise, une évolution de la dynamique des ouvertures / fermetures, etc… Cela peut constituer des indicateurs sur la qualité du dispositif de contrôle interne ou l'évolution de la culture des risques. Mais le point commun à toutes ces illustrations est qu'il faut commencer par avoir des outils de collecte de données robustes et appropriés. L'analyse de données en soi n'est pas forcément la partie la plus compliquée. Agréger, compiler et assurer la qualité des données analysées reste le "nerf de la guerre", au niveau des fonctions de contrôle comme au niveau des métiers, sans oublier l'adhésion des responsables opérationnels dont ce sont les données.



Pour aller plus loin :

- An auditor's guide to Data Analytics
- Les métiers à l'heure de la Data
- Data analytics in the insurance industry
- Big data : l'hôpital soigne ses données médicales

3BIS.

COMPLÉMENT SUR LES OUTILS

TENDANCES ET RETOURS D'EXPÉRIENCE SUR LES OUTILS D'AUDIT ET RISQUES POUR L'UTILISATION DE L'ANALYSE DES DONNÉES (1).

La synthèse ci-dessous est loin d'être exhaustive mais elle permet de voir les différentes tendances en termes d'outils utilisés.

Trois organisations autour des systèmes d'information se dessinent dans les départements audit et risques : le systèmes « core » sur la collecte des données et les satellites qui sont notamment sur l'analyse de données et la visualisation des résultats. Des innovations commencent à apparaître sur des domaines de contrôles tels que le suivi de la réputation cherchant à aller vers de l'intelligence artificielle (IA).

1- Outils core

Les principaux outils utilisés sont : Archer, iCare, Open Pages. Ces outils sont en général pluggés sur des systèmes type ERP (SAP étant le plus utilisé). Ces outils sont utilisés comme référentiels mais la plupart des acteurs y ajoutent des outils d'analyse de données ou de visualisation (MetricsStream, Qlik, Tableau, Spotfire).

Souvent, les solutions retenues dépendent des systèmes utilisés dans les sociétés. Par exemple des problèmes parfois pour connecter avec SAP, implique de remettre des couches applicatives entre les données audit et les données de ces systèmes.

RSA Archer est le plus développé au UK : https://goo.gl/c8FQ9x

2- Analyse de données et visualisation

Les analyses de données sont utilisées plutôt sur fraude, corruption, dépenses des employés et appels d'offres.

Des exemples de couches logicielles sur l'analyse de données :

- https://www.tidytextmining.com/sentiment.html
- https://www.r-project.org/
- https://powerbi.microsoft.com/en-us/
- https://www.operisanalysiskit.com/

Sur les outils de reporting, de plus en plus Excel est remplacé par des outils comme Qlik, Tableau ou Spotfire :

- https://www.qlik.com/fr-fr
- https://www.tableau.com/fr-fr
- https://spotfire.tibco.com/fr

3- Tendances innovantes

Dans les recherches et les innovations basés sur les données, il y a par exemple l'analyse du risque de réputation sur l'analyse des risques de réputation (sur la base des réseaux sociaux) avec Polecast : http://www.polecat.com/

L'utilisation du logiciel IBM watson se développe et est le plus connu. cependant sur les fonctions d'audit il reste utilisé de manière "classique" sur le suivi de recommandations. Watson est orienté sur l'intelligence artificielle, l'outil a du potentiel mais peine à trouver des applications en audit, contrôle interne et gestion des risques : https://www.ibm.com/watson/

Des outils de « process mining », qui permettent de cartographier les processus et suivre leur performance. Complément de l'analyse par les risques. Un exemple avec Celonis. Compliqué à déployer quand les structures sont organisées par risques ou par fonctions. Celonis est souvent utilisé avec SAP parce que les interfaces entre les deux systèmes fonctionnent correctement : https://www.celonis.com/

Conclusion

Les systèmes d'information utilisés par les équipes d'audit et de risques sont en transformation dans leur utilisation et leur fonctionnement. L'objectif est de gagner en efficacité et en couvertures des risques. Les outils de BI ou de big data ont la cote mais sont encore très peu utilisés mis à part sur les sujets de fraude interne.

Dans les difficultés rencontrées, il y a toujours les classiques problèmes de qualité de données, de gouvernance, de reporting avec des couches logicielles qui s'empilent et génèrent des risques opérationnels. Le chemin vers Long et cher pour peu de retours sur investissements à ce stade.

---

(1) Cette analyse est basée sur la réunion « FTSE20 » des responsables d'audit, tout secteur, sur 20 entreprises UK.