

LES LIVRES BLANCS NUAGEO

*Classifier ses
données*

N°4



Conseil Cloud | Nuageo

Sommaire

Avant propos	2
Identifier les besoins en matière de données	3
Identifier les utilisateur et leurs profils	4
Synthèse des besoins	5
Construire des classes de données	6
Conclusion	7



Ce livre blanc est sous licence Creative Commons [CC BY-NC 4.0](https://creativecommons.org/licenses/by-nc/4.0/). Vous êtes libre de le partager, de l'utiliser ou de l'adapter à des fins non commerciales. Pour cela, il vous suffit de citer explicitement son titre, ses auteurs, sa source et la licence à laquelle il se rattache :

« Les livres blancs Nuageo – Classifier ses données » par Nuageo, diffusé sous CC BY-NC

Plusieurs éléments sous licence Creative Commons ont participé à la réalisation de ce livre :

« [Mic](#) » par Shastry; « [Man](#) » par Eugen Balyakoff ; « [Leo Cat](#) », « [Virgo Cat](#) » et « [Taurus Cat](#) » par Denis Sazhin ; « [Megaphone](#) » par NOVITA ASTRI ; « [Target](#) » par Francesco Terzini ; « [data encryption](#) » par Vectors Market ; « [People search](#) » par Eliricon. Accessibles sur thenounproject.com – utilisés sous licence CC BY et modifiés.

Avant-propos

L'objectif de ce livre est de présenter comment mener un projet de classification des données, afin de piloter la stratégie de valorisation de ces données par l'utilisation de nouveaux outils, notamment ceux du Cloud.

L'utilisation des données disponibles dans l'entreprise est le levier principal pour passer véritablement le cap de la transformation numérique. Cependant, cette utilisation des données requiert la maîtrise des données disponibles au sein de l'entreprise. Cette maîtrise passe par la mise en place d'une classification et la définition d'exigences.

Avec les classes de données identifiées, l'entreprise va formaliser le sujet et établir un langage commun pour que les équipes puissent s'en emparer, réfléchir et agir malgré la variété et le nombre de données utilisées.

Dans ce livre, nous vous proposons un ensemble d'éléments à considérer pour votre classification des données.

Cette approche reflète notre vision de la transformation digitale et de l'utilisation de services innovants : d'abord au service des utilisateurs, avec une valeur ajoutée au niveau des usages, tout en apportant flexibilité et confiance.

Vous avez des questions sur le Cloud ? C'est notre spécialité !

Nos livres blancs et autres publications sont à retrouver
sur notre [blog](#)

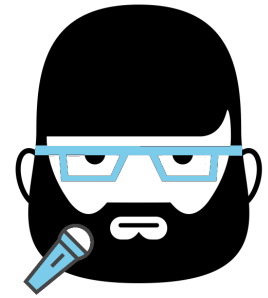


Identifier les besoins en matière de données

La phase préalable à toute classification des données est d'identifier les besoins de chaque équipe ou de chaque service.

Ces besoins doivent ensuite être catégorisés et traduits en niveaux de qualité de service.

Généralement, les catégories utilisées sont les suivantes :



Besoin de disponibilité

Le besoin de disponibilité est lié à la possibilité d'accéder et d'utiliser une donnée ou une application. Il se traduit notamment par le taux de disponibilité, qui représente le temps pendant lequel une donnée ou une application est effectivement accessible sur une période.

L'incapacité de l'entité à travailler en mode dégradé et à conserver sa valeur ajoutée permet notamment d'identifier une exigence forte.

Besoin d'intégrité

Les besoins d'intégrité font référence à l'inviolabilité des données. Il s'agit de garantir le fait que la donnée n'a pas été altérée après son stockage ; ce qui nécessite des contrôles.

Cette exigence peut être incontournable, notamment pour répondre à des exigences réglementaires.

Besoin de confidentialité

La confidentialité des données doit être assurée pour les échanges autour d'éléments sensibles pour l'activité de l'entreprise. Elle se traduit généralement par des niveaux de restriction sur les accès.

Des solutions d'authentification et de chiffrement, ainsi que la restriction de certains accès, permettent d'adresser ce sujet.

Besoin de traçabilité

La traçabilité doit permettre d'identifier qui a saisi, modifié, traité ou accédé à des données. Elle peut se traduire par la mise en place de captures et d'analyses de logs.

Cette exigence est souvent liée à l'exigence d'intégrité pour des raisons réglementaires.

Autres besoins

Ces besoins sont des exigences complémentaires, permettant notamment d'adresser des points relatifs à la réglementation de l'activité de l'entreprise.

Attention cependant à ne pas multiplier les catégories de besoins pour qu'elles restent partageables avec un maximum d'équipes et de services.

On peut penser à des besoins particuliers au niveau de l'archivage (durée, technologie, temps d'accès,...), du chiffrement (technologie, stockage des clés,...), de la réversibilité (facilité de mise en oeuvre, capacité à l'effectuer,...). On peut aussi mentionner les cadres normatifs tels que le respect de SOX pour les paiements, HDS pour les données de santé, ou le RGPD pour le traitement et le stockage des données personnelles.

Identifier les utilisateurs et leurs profils

Une fois que les besoins ont été recueillis auprès des différentes équipes et services de l'entreprise, il faudra prendre en compte deux autres éléments : **la criticité de l'activité** et **le profil des personnes impliquées**.

Criticité de l'activité

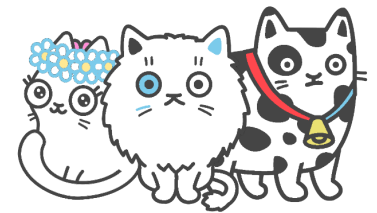
La criticité s'entend ici par la proximité de l'entité étudiée par rapport à la nature de son activité et au cœur de métier de l'entreprise. Il s'agit de prendre en compte l'importance de l'entité pour l'activité de l'entreprise (sa participation au chiffre d'affaires, son importance dans le respect de la réglementation etc.).

Ce critère d'évaluation métier permettra de prioriser les besoins et le planning de déploiement.

Taxonomie des profils

Cette taxonomie doit proposer des types de profils communs à chacune des entités.

Pêle-mêle, on retrouve par exemple des opérationnels, des managers, des prestataires, des stagiaires,...



L'identification de ces profils permet de piloter la gestion des droits d'accès aux données, mais aussi les efforts de sécurité. En effet, la présence de certains profils clés, notamment ceux liés à la direction de l'entreprise, peut impliquer des besoins plus importants que ceux identifiés préalablement. Cela permet donc d'avoir une vision plus juste des besoins et des efforts à réaliser.

On peut retrouver les profils types suivants :

Haut responsable hiérarchique

La notion de responsabilité hiérarchique est déterminante dans la définition des droits d'accès aux données, et dans l'identification du profil de risque de l'utilisateur.

Un utilisateur avec une forte responsabilité hiérarchique est potentiellement plus exposé aux informations sensibles de l'entreprise, et présente donc une attractivité renforcée pour des attaquants. Si ces profils sont présents dans l'activité ou le service, le nombre d'éléments de sécurité à mettre en place peut être plus important.

Opérationnel métier

Quand une personne participe à une activité proche du cœur de métier de l'entreprise, les données qu'elle utilise influencent directement le revenu de l'entreprise ; leur divulgation, leur perte ou leur altération impliquent donc potentiellement une perte financière.

Par conséquent, les opérationnels métier ont vocation à manipuler des données plus sensibles, plus susceptibles d'être la cible d'attaques informatiques ou victime de bugs. Là aussi, le nombre d'éléments de sécurité à mettre en place peut être plus important.


Profil transverse

Au sein de chaque entité il existe des profils qui ont pour mission de fluidifier le fonctionnement de l'équipe, et qui disposent de droits d'accès étendus aux données et applications du service.

Ces profils particuliers doivent être pilotés dans leur consommation des données, notamment pour pouvoir agir en cas de compromission de leurs accès. Ce pilotage demande un effort supplémentaire.

Synthèse des besoins




A l'issu du travail réalisé, l'entreprise doit être en mesure de synthétiser les éléments recueillis pour chaque entité :

Activité de production - Documents relatifs au procédé de fabrication	
Criticité de l'activité	Toute interruption d'activité a un impact direct sur le CA
Profils en présence	
Besoin de disponibilité	24/24, tous les jours ouvrés
Besoin d'intégrité	Les documents ne doivent pas pouvoir être supprimés sans autorisation
Besoin de traçabilité	Il faut pouvoir présenter la liste de toutes les personnes qui se sont connectées au cours du dernier mois si les autorités le demandent
Besoin de confidentialité	Les documents ne doivent être accessibles qu'aux personnes identifiées pendant les heures d'ouverture des locaux

Une fois que cette synthèse a été réalisée pour chaque entité, il sera possible d'identifier des niveaux d'exigences communs à plusieurs d'entre elles. L'objectif est d'obtenir une échelle qui permettent de comparer les besoins et de prioriser les projets.

Avec la criticité de l'activité, l'écart constaté entre les exigences et l'état actuel du système, le décideur aura une vue générale de la situation et pourra choisir les premières études de faisabilité à réaliser selon l'importance qu'il accorde à chaque critère.

Dans l'exemple présenté ci-dessous, l'intégrité et la traçabilité des données employées par les activités d'approvisionnement et de production sont les deux axes d'amélioration les plus importants. Cependant, du fait de la criticité de l'activité de production, celle-ci pourra être priorisée.

	Production			Vente			Ressources Humaines		
Criticité	Très Forte			Forte			Moyenne		
Profils									
Etat	Actuel	Attendu	Ecart	Actuel	Attendu	Ecart	Actuel	Attendu	Ecart
Disponibilité	Fort	Très Fort	+1	Fort	Très Fort	+1	Fort	Moyen	0
Intégrité	Moyen	Très Fort	+2	Faible	Fort	+3	Faible	Faible	0
Traçabilité	Moyen	Très Fort	+2	Moyen	Fort	+2	Faible	Moyen	+1

Confidentialité	Fort	Très Fort	+1	Fort	Très Fort	+1	Fort	Très fort	+1
-----------------	------	-----------	----	------	-----------	----	------	-----------	----

Construire des classes de données

Les analyses mentionnées auparavant permettent de faire émerger la classification des données. Cette dernière doit reprendre les besoins identifiés lors des entretiens réalisés avec les équipes, mais elle doit proposer des classes qui les regroupent et correspondent à des usages partagés par tous les services. L'objectif est de donner un outil compréhensible et utilisable par les équipes impliquées dans la valorisation de leurs données.

Cette classification va permettre aux équipes de classer elles mêmes leurs données et d'adopter les usages adéquats à chaque classe de données pour répondre aux besoins identifiés.

A titre d'exemple, nous vous proposons la classification suivante :



N1 - Publique

Les données publiques sont celles qui sont typiquement utilisées et publiées sur les sites institutionnels. Les besoins sont donc limités car par nature ces données présentent un caractère peu sensible et peu critique pour l'activité.



N2 - Interne

Il s'agit de données le plus communément utilisées au sein de l'entreprise. Ce sont des données habituellement manipulées par tout employé dans le cadre de ses activités, et qui n'ont pas vocation à être partagée naturellement au public.



N3 - Confidentiel

Ces données présentent un caractère plus sensible que les données internes, et sont relatives au coeur de métier de l'entreprise.

Ces données peuvent par exemple révéler des processus internes, des secrets industriels, ou des informations sensibles quant à l'approvisionnement ou la tarification.



N4 - Données personnelles

Avec l'entrée en vigueur du Règlement Général sur la Protection des Données personnelles, un effort particulier doit être apporté quant aux traitements effectués sur des données personnelles.

Ces données, ainsi que les processus liés à ces données, doivent être clairement identifiés et faire l'objet d'une approche spécifique.

	1. Publique	2. Interne	3. Confidentiel	4. Donnée personnelle
Disponibilité	Très fort	Fort	Très fort	Moyen
Intégrité	Fort	Fort	Fort	Très fort
Confidentialité	Faible	Fort	Fort	Très fort
Traçabilité	Faible	Fort	Fort	Fort

Conclusion

Un projet de classification des données s'inscrit dans une approche plus générale de transformation et de modernisation des activités de l'entreprise.

La migration vers des usages Cloud est notamment un cas dont l'une des étapes préliminaires est l'identification des données qui vont être consommées de cette manière : la classification s'inscrit donc dans une démarche de maîtrise de ces nouveaux usages, afin d'accompagner les utilisateurs vers une utilisation raisonnée et pertinente de ces outils.

Ce chantier est impressionnant : il demande l'implication de l'ensemble des métiers de l'entreprise. Il ne peut pas être porté uniquement par les équipes IT, malgré leur connaissance et maîtrise du patrimoine des données.

A ce titre, la classification des données est un point incontournable pour tout projet de transformation : la donnée s'est imposée comme une ressource indispensable pour renforcer l'activité de l'entreprise, et l'exploitation de ces données doit s'effectuer avec maîtrise et confiance.

Ces éléments se construisent avec des outils, et surtout avec la concertation des utilisateurs. Ils sont la pierre angulaire de tout projet de transformation et doivent être au coeur de ces projets.



Société de conseil spécialisée dans la transformation numérique des entreprises

Nuageo.fr