

Nouvelles efficacités et création de valeur : les projets de gouvernance des données

Ce cahier porte sur la gouvernance des données qu'elle soit mise en place en inter-services ou en multipartenariat. A partir de nombreux exemples et de plusieurs préconisations, il s'efforce de présenter un panorama des particularités issues d'une gouvernance des données.

Ce cahier est le résultat d'un travail et d'échanges collectifs réalisés tout au long de l'année 2016 au sein de la Campagne Infolab de la FING. Le groupe a été animé par Armelle Gilliard, avec l'étroite collaboration de Charles Népote et la complicité de Sarah Labelle et Simon Chignard. Le groupe était composé d'Amandine Crambes, Sarah Marquet, Valérie Weber-Haddad, Martin Régner pour l'Ademe, Axelle Reynier pour le Groupe UP, Marine Guillon, François Herlent pour la Maif, Isabelle Micheu pour La Poste et Thomas Bekkers pour la Région PACA. Le cahier a été rédigé sous la direction d'Armelle Gilliard.

Retrouver ce document en ligne :

<http://infolabs.io/gouv16>

1

La gouvernance des données, de quoi parle-t-on ?

Dans son rapport de 2015 sur la “gouvernance de données”, l’administrateur général des données de l’Etat (SGMAP), la définit comme “l’ensemble de principes et de pratiques qui visent à assurer la meilleure exploitation du potentiel des données (p. 48).” Le terme est générique, se référant à des principes, donc un cadre, mais aussi des pratiques.

Le terme de gouvernance est apparu dans les années quatre-vingt dix. Comme la notion de gouvernement, il reprend l’idée d’organisation de la cité et de structure. Mais la gouvernance associe la dimension institutionnelle à une dimension participative et dynamique dans la gestion des affaires publiques. Dans les faits, elle suppose un partage de la prise de décision entre membres de la structure qui ont souvent des intérêts divergents. Ainsi, la gouvernance est une forme de pilotage partagé entre les (multi-)parties prenantes dans une organisation. La neutralité du Net est un principe qui repose sur la philosophie de la gouvernance. Quand on parle de gouvernance, on travaille à la responsabilisation des acteurs lors des prises de décision, en s’intéressant à la diversité des acteurs et des modes de participation possibles.

La gouvernance peut se matérialiser par une charte, document peu contraignant qui décrit la façon dont la gouvernance fonctionne. Son existence n’est pas indispensable. Elle est plutôt révélatrice d’un état de maturité de la gouvernance. Elle rassemble les règles sur lesquelles tous les groupes sont d’accord.

Notre sujet est la gouvernance des données : il s’agit de la gouvernance qui permet une ges-

tion optimale des données dans la ou les structures. Ce document souhaite aider à mettre en place une gouvernance des données ou à en faire son suivi. Il ne couvre pas la totalité de ce très vaste sujet. Il est organisé en deux parties : une partie théorique pour comprendre et une partie pratique sous forme de questionnaire qui vous aidera à évaluer votre gouvernance des données et vous facilitera la mise en place d’actions de votre démarche.

2

Impulser la démarche : quatre catalyseurs clés

Le premier enseignement de notre exploration est que la gouvernance intervient lorsqu’un problème dépasse le périmètre d’action de chacun des individus ou des groupes. Le besoin de gouvernance naît de l’implication de multiples acteurs. Bâtir une gouvernance, qu’elle soit au sein de l’entreprise, avec des partenaires ou dans des démarches de co-constructions, va constituer autant d’opportunités de valoriser les compétences de chacun. Une vision stratégique sur l’usage auquel répondent les données est essentielle pour une gouvernance efficace. C’est pourquoi il sera très important d’identifier précisément ce qui cimente le groupe en charge de la gouvernance. Nous pensons que pour qu’une gouvernance fonctionne bien, elle doit s’articuler autour d’au moins un catalyseur.

Le catalyseur est ce qui va motiver et aider les personnes à travailler ensemble. Le fait d’identifier le catalyseur principal va clarifier les objectifs pour tous les acteurs. Nous avons identifié quatre catalyseurs distincts.

2.1 L'optimisation de ressources

Les acteurs recherchent une forme d'économie, d'optimisation des ressources. La gouvernance va permettre de limiter les coûts, mutualiser les ressources et/ou de partager les dépenses.

Trois exemples de projets avec une gouvernance des ressources

1. Depuis 20 ans, le GIP ATGeRi au travers PIGMA (Plateforme d'Information Géographique Mutualisée en Aquitaine - <https://www.pigma.org/>) gère un patrimoine de données de 650 partenaires (soit plus de 6500 informations cataloguées, partagées et échangeables) qui est à la disposition des décideurs publics aquitains. L'origine du projet s'est fondée sur des acquisitions de fonds cartographiques mutualisés.

2. Le projet Optimod Lyon (<http://www.optimodlyon.com/>) est organisé et coordonné par le Grand Lyon. Il réunit 13 partenaires publics et privés pour construire une plateforme intégrée d'innovation sur la mobilité urbaine et accélérer la coopération public-privé sur les systèmes de transports intelligents en milieu urbain.

3. La Régie de Gestion des Données des Pays de Savoie (RGD 73-74 : <http://www.rgd73-74.fr/>), est une structure conçue comme un service public bi-départemental pour les collectivités locales et les organismes publics, sous le statut de régie à caractère industriel et commercial. C'est un outil de connaissance du territoire à disposition des collectivités, pour les aider à répondre aux enjeux territoriaux en se basant sur des éléments objectifs. Une base de données agricoles a été constituée à moindre coût et un Observatoire des Territoires de la Savoie a vu le jour.

>> Pour cette opportunité, il sera important de mettre en place un **accord de partenariat**. Il aura pour points clés :

- la forme du partenariat ;
- les droits d'accès ;
- le prix du service ;
- la durée du service.

2.2 La création de valeur pour les acteurs de la gouvernance

Les acteurs ont des objectifs d'entreprise différents. A travers les compétences de chacun, la gouvernance apporte des complémentarités pour créer de la valeur. Cette forme de gouvernance bénéficie directement aux acteurs qui adhèrent au projet.

Trois exemples sur la gouvernance créatrice de valeurs

1. La BAN (Base Adresse Nationale - <https://adresse.data.gouv.fr/>) regroupe 3 acteurs aux complémentarités bien pensées. La Poste est forte sur l'adresse, l'IGN sur la précision technique et géographique et OSM porte des qualités d'animation et le point de vue des réutilisateurs. Ainsi les opportunités pour chacun des 3 partenaires se répartissent comme suit :
La Poste : opportunité économique et recherche d'efficience,
IGN : opportunité d'efficience et de notoriété, compétences en informatique géographique
OSM : opportunité d'efficience, compétences internationales et en gestion de crises

2. Le projet Data littoral (www.littoral-aquitain.fr) regroupe plusieurs acteurs du littoral aquitain pour trouver des solutions concernant la migration massive et ponctuelle de la population de la métro-

pole vers la côte Atlantique. De ce groupe de travail hétérogène sort de nouveaux groupes de travail sur le tourisme et la protection de la nature.

3. Dans le projet Pigma (voir plus haut), les données géographiques sont partagées afin de surmonter la crise occasionnée par la tempête Klaus dans le Sud-Ouest (2009). Différentes actions sont mises en place dont une cartographie <http://cartographie.environnement-aquitaine.fr/visualiseur/?idlyr=11397>

Pour cette opportunité, il sera important de mettre en place des actions de co-production. Elles auront pour points clés :

- l'identification des compétences de chacun ;
- l'identification du rôle de chacun ;
- l'identification des ressources de chacun ;
- avoir des outils bien adaptés pour travailler ensemble et bien organiser le travail et les travaux communs ;
- co-produire le cahier des charges sur un espace partagé ;
- disposer d'un animateur, facilitateur et valorisateur de la co-production (rôle de chef d'orchestre).

2.3 La création de valeur favorable à un écosystème

La gouvernance crée de la valeur pour un écosystème qui concerne les acteurs de la gouvernance et au-delà. Cette forme de gouvernance a un effet de levier, elle profite à des bénéficiaires qui ne sont pas les acteurs de la gouvernance.

Trois exemples de gouvernance au service d'un écosystème :

1. Le système AIS pour les bateaux qui peut être illustré par le service de Marine Traffic (<http://www.marinetraffic.com/>). Ce service affiche en temps réel les mouvements des bateaux dotés d'émetteurs AIS. Il est obligatoire pour les navires de jauge supérieure à 300 tonnes mais de nombreux bateaux, plus modestes, en sont équipés, du fait qu'il permet d'éviter les collisions. 1881 stations référencées suffisent à couvrir la quasi intégralité du trafic maritime mondial. Il y a 82 stations en France. Le capteur installé à Marseille couvre une zone de 3200 Km² (jusqu'en Algérie). Au même titre que le GPS, c'est un catalyseur de création de valeur qui touche un écosystème.

2. Transfermurga (www.transfermurga.eu/fr/) est un projet initié en 2014. Suite à une analyse approfondie du transport transfrontalier France-Espagne, différentes actions ont été envisagées et notamment Transfermurga (<http://www.transfermurga.eu/fr/>) un portail, basque, espagnol et français, d'information sur l'offre de transports, publics et privés, transfrontalière propulsé par l'Eurorégion Aquitaine-Euskadi. Le portail possède la fonctionnalité de calculateur d'itinéraire multimodal. Le projet inclut la promotion de l'open data et favorise l'innovation.

3. Le programme DatAct (<http://www.dataact.fr/>) a pour objectif d'amener les acteurs des collectivités à imaginer de nouvelles formes d'action publique et de mise à disposition d'outils basés sur le partage des données dans des nouveaux modèles de coopération. Le programme Dataact vise à modéliser une Régie de données afin de faire qu'une économie de la donnée soit viable pour tous les acteurs parties prenantes de la production de services territoriaux.

Pour cette opportunité, la stratégie partenariale à mettre en place sera de type gestion de biens communs.

Définition de la notion de biens communs (Elinor Ostrom)

« Le terme de ressources communes désigne un système de ressources suffisamment important pour qu'il soit coûteux (mais pas impossible) d'exclure ses bénéficiaires potentiels à l'accès aux bénéfices liés à son utilisation. »

Ex l'usage d'un terrain communal pour les animaux d'une personne.

La stabilité de la gestion de biens communs se construit par « petits pas » qui sont bénéfiques aux acteurs. Les 8 règles minimum pour que l'organisation soit stable sont :

- Existence de limites clairement définies :
 - sur les individus ayant accès aux ressources,
 - sur les limites de la ressource ;
- Adaptation aux conditions locales (moyens matériels, financiers, humains) ;
- Existence de choix collectifs incluant la plupart des individus concernés ;
- Existence de modalités de surveillance des comportements des individus ayant accès à la ressource ;
- Existence de sanctions graduelles en direction des individus qui transgressent les règles ;
- Existence de mécanismes de régulations des conflits rapides et bon marché ;
- Reconnaissance minimale par les autorités externes du droit à l'auto-organisation ;
- Imbrication des institutions locales au sein d'institutions de plus grandes échelles.

2.4 L'application de la réglementation

La gouvernance donne le cadre pour mettre en place une contrainte juridique des membres qui la composent. Cette forme de gouvernance est plus pilotée par les attendus de la loi que par ces acteurs.

Des obligations réglementaires peuvent à elles seules justifier la mise en place d'une gouvernance des données : par exemple, le règlement européen sur la protection des données voté en avril 2016¹ et applicable pour nombre de structures dès mai 2018 (voir parmi les exemples). De plus, certaines professions sont soumises à une réglementation particulière comme les activités d'intermédiation en assurance ou dans le secteur alimentaire (traçabilité...).

Les avantages d'un catalyseur réglementaire consistent dans le suivi de qualité des données.

Cinq exemples de gouvernance en lien avec la réglementation :

1. Le règlement européen du 27 avril 2016 sur la protection des données pose le principe général de responsabilité du responsable de traitements. Cela se traduit par une approche par les risques - le responsable de traitement doit recenser et auditer les traitements et adapter sa politique de protection des données au niveau de risques. Le responsable de traitements doit notifier, dans les meilleurs délais (maximum 72 heures, sinon il faut justifier) de toutes failles de données à caractère personnel. Le principe de protection des données se fait dès la conception (privacy by design) : proportionnalité, conservation, accessibilité, accès, analyse d'impacts, portabilité...

1. Règlement (UE) 2016/679 : <http://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A32016R0679>

2. Les Infrastructures de Données Géographiques nationales et régionales s'inscrivent dans l'effort (guidé par la directive européenne INSPIRE) de standardisation des données et des modèles par de nouvelles règles de mise en oeuvre. En particulier, ces règles incluent des modalités de géoréférencement, la définition et classification des objets géographiques et des spécifications sur la structuration des futures bases de données. Le but de cette standardisation est de faciliter l'interopérabilité des données et des systèmes

3. Le décret n°2016-972/93 du 18 juillet 2016 fixe l'obligation d'ouverture et de mise à disposition des données de production et de consommation énergies des opérateurs, dans le cadre de la loi sur la croissance verte et la transition énergétique. Les données seront récoltées par le Service de l'observation et des statistiques (SOeS) entre 2016 et 2020. L'intérêt d'opérer une ouverture ou une mise à disposition à l'échelle d'un territoire a beaucoup de sens. La région Provence-Alpes-Côte d'Azur expérimentera prochainement, avec les énergéticiens, une gouvernance partagée portant sur l'agrégation des données Energie concernées par les décrets mentionnés ci-dessus.

4. Respect des secrets légaux

Les entreprises sont concernées par plus ou moins de secrets légaux selon leur(s) secteur(s) d'activité(s). Par exemple, le secret statistique (loi du 7 juin 1951) permet d'assurer :

- aux personnes physiques que la confidentialité sur leur vie personnelle et familiale sera garantie ;
- aux entreprises que le secret commercial sera respecté.

Autre exemple : les données relatives à la santé des personnes, et plus généralement les données sensibles, bénéficient d'une protection spécifique et sont non communicables, sauf aux personnes qu'elles concernent ou sur la demande d'un médecin.

“Le plus important, dans le cas d'un secret légal, est donc de bien savoir à qui il s'oppose, et dans quelles conditions.” Rapport 2015 de l'administrateur général des données.

5. Les assureurs et les banquiers ont pour obligation réglementaire européenne de rendre compte à l'ACPR (Autorité de Contrôle Prudentiel et Réglementaire). Cela permet de limiter les risques de solvabilité en prouvant que le calcul de la solvabilité est correctement fait.

Pour cette opportunité, la stratégie partenariale à mettre en place sera de type transformer une contrainte en opportunités. L'avantage de ce catalyseur est que tout le monde doit aller dans le même sens. Dans la mesure du possible, il est intéressant de le combiner avec une autre opportunité pour en diminuer la dimension contrainte.

2.5 Comment s'appuyer sur les catalyseurs ?

Le catalyseur doit être bien défini et précis pour inciter les partenaires à venir aux réunions, partager des données et travailler ensemble.

Il doit répondre aux questions suivantes :

- Qu'est-ce qui fait qu'on a besoin d'être autour de la table ?
- Qu'est-ce que ça va apporter de se fédérer ?
- Quelles sont les opportunités/avantages que l'on va tirer du processus ?
- Quels sont les objectifs communs ?
- En quoi ce qu'on va faire va produire ce qu'on en attend ?

Il peut y avoir un ou plusieurs catalyseurs pour

un même groupe d'acteurs. Plus il y a de catalyseurs, plus les liens et l'implication des acteurs seront forts (cautionnés par la hiérarchie de chacun). Un catalyseur peut, avec l'évolution du projet, s'enrichir d'autres catalyseurs.

Le catalyseur peut être conçu de manière macro sur un objectif important (sur 3 séances de travail ou plusieurs années), mais il est intéressant de l'aborder sous un angle micro : un tour de table qui permet aux partenaires de présenter les données qu'ils ont chez eux et qu'ils peuvent partager.

3

Cartographier

Dans bien des cas, dans l'organisation, les données ont été produites avec l'émergence de l'informatique, au fil des besoins exprimés par les différents collaborateurs. Elles sont une production évidente de l'entreprise. Peu à peu, l'organisation utilise ses données en dehors de leurs productions d'origine, ce stade s'apparente à la cueillette.

Si l'on veut aller plus loin dans la gouvernance des données, il faudra passer par une approche plus structurée, se mettre en posture de gestionnaire des données de l'organisation. Cette conquête du territoire numérique de l'entreprise sera plus facile à mener si l'on passe par un travail cartographique. Sans ce travail, il ne sera pas possible de prendre conscience de toute l'étendue du sujet et de prioriser les actions. Sans cette forme de représentation, il sera impossible de transmettre aux personnes dont la culture numérique est faible, ce qu'on est en train de faire et pourquoi c'est important pour l'entreprise de le faire.

Nous vous recommandons deux formes de cartographies :

- une qui identifie les points forts et faibles des compétences de l'entreprise et les politiques qui doivent être mises en place pour traiter les risques ;
- une qui identifie les données dont a besoin l'organisation, qu'elle les possède ou non.

Ce travail de cartographie aide à construire une gestion patrimoniale des données, qui pourra être découpée, phasée, priorisée, budgétée.

Dans le cas d'une gouvernance multi-partenaire, le travail de cartographie doit se faire au niveau de chaque partenaire et être partagée au niveau du groupe. Chacun doit avoir une vue transversale et cohérente de ce qu'il apprend et de ce qu'il apporte.

3.1 Cartographier les expertises des données dans l'organisation

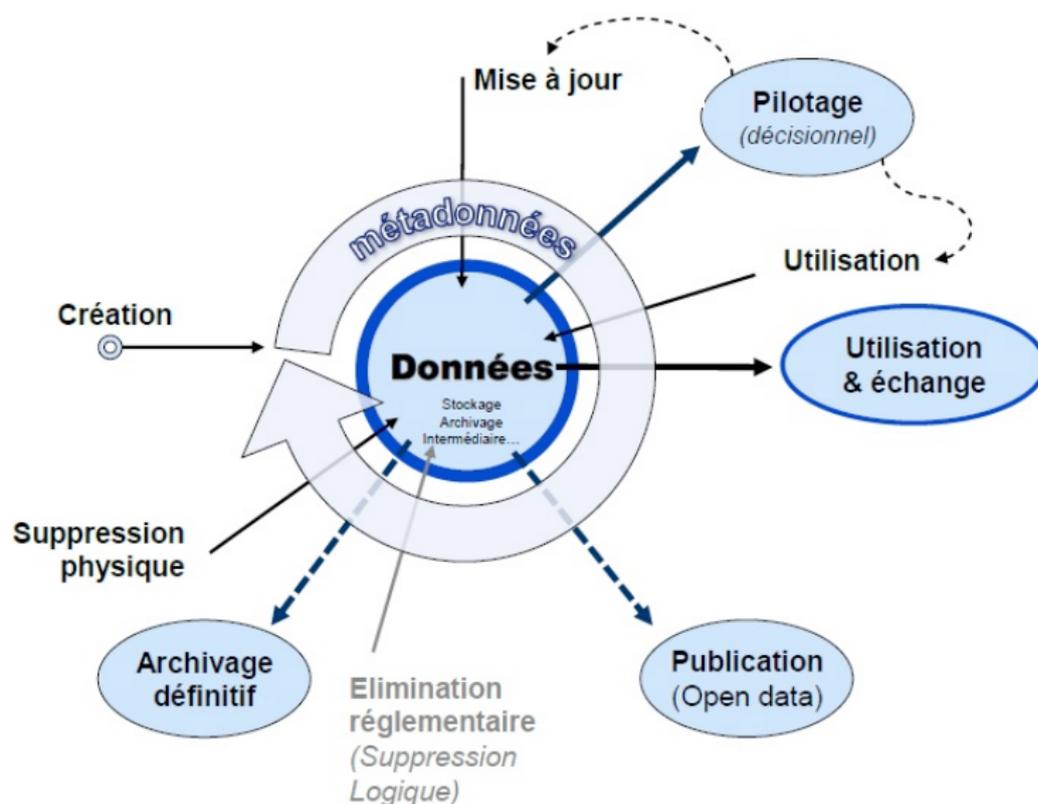
La cartographie des expertises des données au sein d'une organisation est l'une des tâches d'un administrateur général des données (chief data officer) dans l'organisation. Début 2017, il est encore assez rare qu'une telle fonction soit formalisée. Quand les compétences existent dans l'entreprise, elles sont le plus souvent dispersées dans différents services : systèmes d'information (y compris géographiques) et de documentation, service juridique, statistique, innovation, communication multimédia.

Quoi

Dans l'organisation, il faut identifier les interlocuteurs, leurs compétences, leurs rôles : s'assurer que le rôle de chacun est bien connu ; s'assurer que l'intégralité du cycle de vie est pris en charge dans l'organisation.

ILLUSTRATION

Cadre Commun d'Urbanisation du Système d'Information de l'Etat (Version : 1.0 du 26/10/2012)



En reprenant le cycle de vie des données (voir schéma), nous avons identifié plusieurs rôles concernés par la gouvernance des données, dès l'amont de leur création (avec l'urbanisation des systèmes d'information) à leur exploitation (qui concerne la valorisation, le suivi des contributions) et leur fin de vie (archivage, suppression physique). Ces éléments sont détaillés en annexe pour que vous puissiez appliquer cette approche à votre organisation.

Il est indispensable de repérer qui s'occupe de quoi dans l'organisation et de s'assurer que le périmètre et les complémentarités de chacun soient précisés. Ce travail doit être fait avec les différents services de l'organisation. Notamment, cela peut consister à demander à chaque

salarié de se situer sur cette cartographie : cela permettra de faire ressortir le besoin de clarifier les rôles en cas de superposition des périmètres individuels, ou au contraire d'identifier des domaines où aucune responsabilité n'a été identifiée.

Le travail de gouvernance portera sur toutes les politiques transversales à ce cycle de vie des données. Citons-en les principaux éléments :

- formation en matière de données : gestion, qualité, usages, etc. ;
- définition de l'organisation : enjeux, fonctions, métiers, profils, postes, etc. ;

- clarification des responsabilités et des règles d'accès aux données ;
- respect des obligations juridiques (dimension compliance) ;
- politique de sécurité des données : politique d'accès, politique de protection, etc. ;
- suivi du cycle de vie des données (stockage, exposition, mise à disposition, archivage) et de leur qualité ;
- suivi des expérimentations et des actions de valorisation des données.

La dimension métier ne doit pas être oubliée, et elle sera bien évidemment très différente selon le secteur d'activités concerné (santé, assurance, commerce, agriculture...).

C'est à l'administrateur des données dans l'organisation de s'assurer que ces différentes politiques soient en place. Cette fonction peut toutefois être portée de manière collective, d'autant plus que les compétences requises sur chacune des thématiques identifiées peut être très spécifiques. Ainsi, le rôle de l'administrateur général des données s'apparente à celui d'un chef d'orchestre, s'assurant qu'aucune dimension de la gouvernance n'est délibérément ignorée.

3.2 Cartographier les données de l'entreprise

Que cartographier ?

Représenter les données de l'entreprise, permet d'en faire un inventaire, et de connaître les liens qu'elles peuvent avoir entre elles. Savoir quel service se sert de quelles données donnera une vue partagée à chacun et confèrera du sens à ceux qui les collectent, à ceux qui les transforment, à ceux qui les publient. Cette simple visualisation partagée peut stimuler les équipes à aller vers une meilleure qualité des données.

Cette représentation des données de l'entreprise sera aussi l'occasion d'évaluer les jeux de données qui constituent la colonne dorsale de

l'activité de l'entreprise et ceux qui sont les plus réutilisés par les acteurs internes et les partenaires.

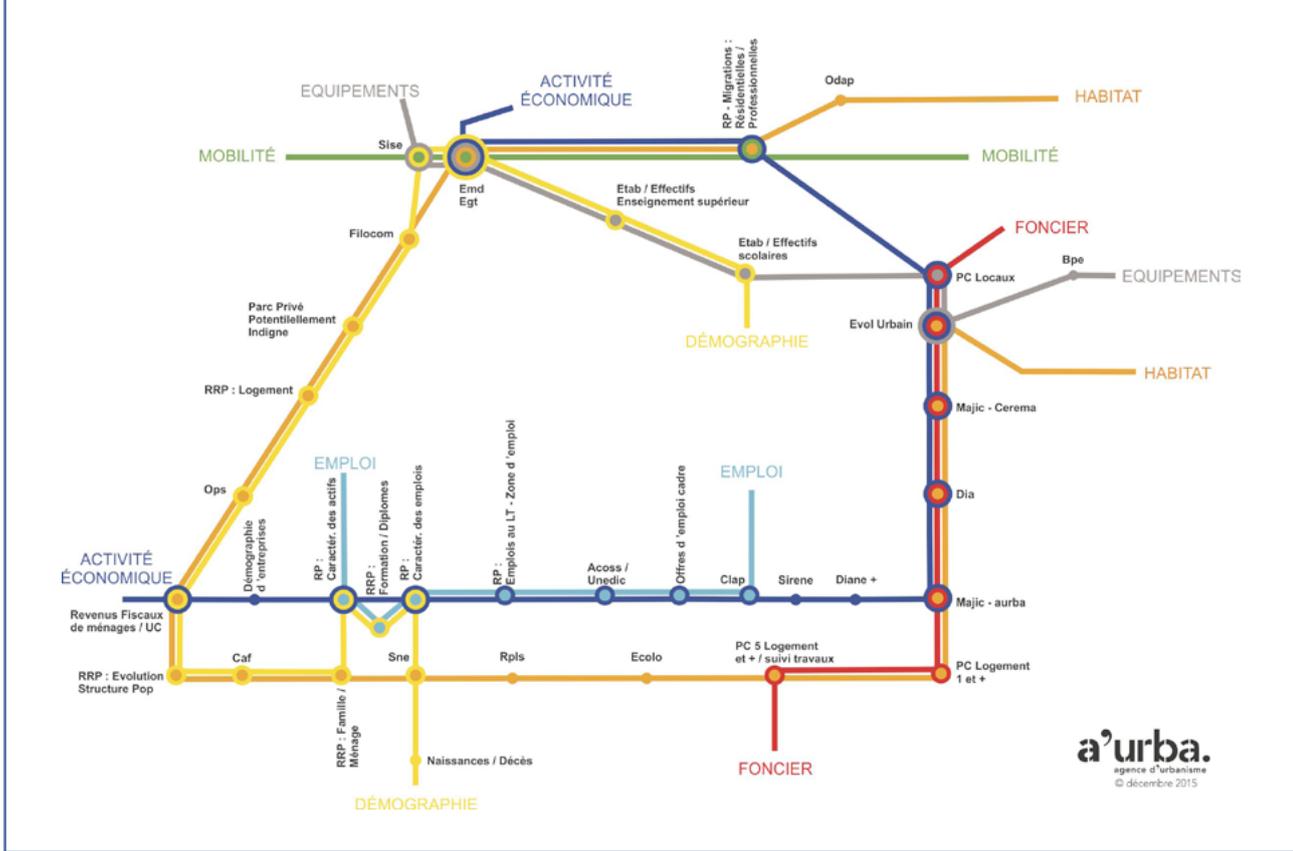
Pour ces jeux de données essentiels à la vie de l'entreprise, un travail sur le cycle de vie de la donnée, sa consolidation, la qualité des données, sera l'occasion de mieux gérer les priorités sur les différentes politiques des données que mène l'organisation (cf. plus haut : cartographie des expertises de données dans l'organisation).

Comment travailler à cette cartographie ?

Il y a plusieurs formes de représentations possibles des jeux de données, nous vous conseillons d'opter pour celle qui vous semble la plus simple et la plus rapide à mettre en oeuvre. A titre d'exemple, l'A-urba (agence d'urbanisme Bordeaux métropole Aquitaine) a élaboré une solution inspirante, sous la forme d'une représentation type "plan de métro". Les lignes représentent les grands thèmes de travail de l'agence et chaque "station" un jeu de données. Cette représentation facilite l'identification des données centrales et met en valeur les thèmes consommateurs de données.

ILLUSTRATION

"plan de métro" A-urba (agence d'urbanisme Bordeaux métropole Aquitaine)



4

Animer la gouvernance des données

L'animation retenue doit être l'occasion de rendez-vous réguliers avec les partenaires et des moyens spécifiques doivent y être consacrés.

L'animation de la gouvernance cherche à mettre les partenaires dans une posture d'amélioration continue et peut s'appuyer sur des questionnaires d'évaluation, des entretiens individuels post-animations. Les modalités de l'animation peuvent être renouvelées régulièrement.

Une des préoccupations doit être de surprendre positivement les partenaires :

- en les encourageant sur les efforts (compliments, participation à des concours/récompenses...);
- en apportant des saisonnalités aux animations ;
- en apportant des éléments humoristiques ou/et ludiques ou/et dépayésants ou/et conviviaux ;
- en changeant de lieux de rencontres ;
- en faisant intervenir un prestataire, un(e) expert(e).

L'offre de l'animation s'organise autant autour de temps d'efforts pour co-construire que de temps d'enrichissements pour les participants. Cela peut se traduire par la formation mutualisée, le partage d'expériences, les rencontres entre pairs, les bons plans (un appel à projet, des sources de financements, des opportunités, des achats groupés, des prix avantageux, des informations pratiques, l'accès privilégié à des services...).

Pour une grande communauté, s'il est nécessaire de dupliquer les animations, les animations sont les mêmes pour tous. Les animations sont pensées dans un programme global de type annuel. Les dates des animations sont annoncées longtemps à l'avance (2 mois). Les invitations sont formelles et anticipées: ordre du jour, horaires, lieux, dates.

L'animation est réussie si vos partenaires deviennent vos ambassadeurs.

Pour les réunions au téléphone, il peut être très utile de partager un document en ligne ou bien d'utiliser un tableau blanc en ligne du type Framapad (<https://framapad.org/>).

Aller plus loin

Travaux du groupe de travail Infolab organisation 2016 sur la culture des données : <http://infolabs.io/cd2016>.

4.1 Expérimenter

Qu'expérimenter ?

L'expérimentation doit porter sur « Qu'est ce qui va servir au groupe de travail ? » le plus rapidement possible. C'est une des manières de donner corps à la gouvernance, de la faire grandir, en plongeant véritablement les mains dans le cambouis. C'est une manière de s'approprier la matière nouvelle créée par la gouvernance. C'est aussi une façon de rendre les données plus tangibles, de les triturer, de produire des croisements imprévus et féconds, rendus possibles par un partage bien gouverné des données.

Comment expérimenter ?

Le retour sur investissement doit être d'autant plus rapide que la gouvernance est jeune. Les premières expérimentations et les premiers résultats sont une première forme d'outillage

mais aussi une opportunité de confirmer l'utilité du groupe.

Les premières expérimentations mettront l'accent sur la complémentarité des compétences et des données nouvellement partagées, créant ainsi une valorisation et un engagement de chaque acteur.

Aller plus loin

Travaux du groupe de travail Infolab organisation 2016 sur les expérimentation avec les données : <http://infolabs.io/xp16>.

4.2 Améliorer la qualité des données

Quel rôle pour la qualité ?

Sujet difficile, voire rébarbatif, serait-il trop éloigné de nos questions de gouvernance ? L'amélioration de la qualité possède pourtant de grandes vertus. Elle favorise la réutilisation des données, elle est un vecteur de dialogue, elle permet aux différentes compétences de s'exprimer et de valoriser leurs complémentarités. Elle touche vraiment tous les acteurs.

Comment favoriser la qualité ?

L'industrie a montré depuis longtemps que la qualité n'est pas seulement un sujet technique, mais un véritable sujet organisationnel. Ce sont les hommes et l'organisation qui déterminent la qualité, et pas les outils. Il n'existe pas de réponse unique et magique à ce problème. Pour autant, les entretiens réalisés formulent quelques bonnes pratiques :

- Donner à voir le cycle de vie des données : il circonscrit le périmètre de production et d'usage, identifie les maillons de la chaîne.
- Plonger rapidement les mains dans les données, par petits groupes hétérogènes: cela crée du dialogue et permet d'identifier les problèmes aigus ou récurrents.
- Privilégier les bonnes pratiques de remontée et de correction des données.
- Veiller à la qualité des métadonnées et de

la documentation, avant même la qualité des données ! (Un point d'intérêt géolocalisé à 300 m près, ce n'est pas forcément de la non qualité.).

Par dessus tout, la qualité est un domaine qui s'anime et qui ne va pas de soi. La gouvernance doit s'en saisir.

Aller plus loin

Le groupe de travail Infolab organisation sur la qualité des données a produit collectivement deux documents :

- une méthodologie d'amélioration de la qualité d'un jeu de donnée, inspiré des méthodes agiles : le sprint qualité : <http://infolabs.io/sprint-qualite>
- le démarrage de la rédaction collaborative d'un véritable guide de la qualité, dont l'écriture se poursuivra en 2017 et au-delà : <http://infolabs.io/qualite-donnees>

4.3 Valoriser l'animation

Quel intérêt de l'animation ?

Les animations constituent la face visible du travail de gouvernance. En faire un bilan et co-produire un document de synthèse qui peut circuler auprès des participants et de leurs responsables hiérarchiques sera un élément clé de la valorisation.

Il doit apparaître comme un retour sur investissement positif au regard du temps et du budget consommés.

Comment valoriser ?

Toute action d'animation est documentée, c'est-à-dire décrite précisément et méthodologiquement. Cela offre la possibilité de communiquer les éléments essentiels si le partenaire a été absent ou s'il veut revenir sur l'animation.

Le bilan des animations peut s'inspirer des qualités d'un communiqué de presse, et être visuel et synthétique. Une forme possible est une data-visualisation de type poster. Il pourra être utilisé

comme illustration d'un article, poster dans un bureau, insérer comme diapositive d'une présentation, placer dans une exposition....

Dans l'idéal, il est un élément d'une série. A terme, il aura une fonction pédagogique et il permettra de se rappeler les différentes étapes de construction de la gouvernance.

Si les étapes suivantes sont relativement maîtrisées, le bilan est une opportunité de les présenter.

Le bilan synthétise « les belles histoires » du projet data. A titre d'illustration, on peut se reporter :

- Aux actions d'acculturation dans l'organisation du cahier de la Fing sur la Culture des données : <http://infolabs.io/dc16>
- La liste des réutilisations de données, établie par les administrations qui ont ouvert leurs données, dans l'annexe 7 du rapport remis par l'Association Opendata France à Estelle Grelier et Axelle Lemaire².

[2. Rapport sur les dispositifs d'accompagnement des collectivités locales à l'ouverture des données publiques \(oct 2016 \) https://frama.link/rapportODF_oct16](https://frama.link/rapportODF_oct16)

5

Documenter la gouvernance des données

Dans l'idéal, on s'appuiera sur une documentation la plus légère possible, utile aux nouveaux entrants et sur laquelle peuvent se référer l'ensemble des participants en cas de doute. Nous avons retenu un exemple de documentation : la charte d'usage des données.

5.1 Exemple : la charte d'usage des données

Pourquoi une charte ?

Cette charte, destinée aux collaborateurs, peut venir compléter la charte des usages informatiques, elle vient préciser cette dernière sur toute la thématique des données.

Comment rédiger une charte ?

Les sujets clés de la charte sont le traitement des données personnelles, les modalités d'échanges et de réutilisation des données, la sécurité des données, les procédures d'informations lorsqu'un problème survient. Elle peut être une synthèse des différentes politiques mises en place concernant les données (voir le paragraphe 3.1 Cartographier les expertises des données dans l'organisation).

L'avantage de cette forme de culture des données est de favoriser une implication des ressources humaines. A sa sortie, il fait l'objet d'une forte communication interne. Le document peut être remis à jour tous les ans et bénéficier, ainsi d'une communication récurrente.

Un tel document peut offrir à chaque organisation les moyens d'aller vers une meilleure maîtrise des risques liés aux données.

Il est possible de donner une orientation marketing à ce document : l'organisation peut s'engager auprès de ses clients sur les différents traitements de données. A titre d'exemple, l'entreprise peut informer ses clients qu'elle fait le choix de ne pas vendre leurs données. Une telle décision devra se traduire au minimum par une procédure interne qui responsabilise les référents sur la vente des données et la justification de tels choix.

Comme on le voit dans cet exemple, la charte est un très bon vecteur d'une culture partagée. Elle va aider à rendre cohérentes les décisions prises dans l'entreprise. C'est au responsable de la gouvernance de la mettre en place et d'assurer cette cohérence.

Conclusion

Une bonne gouvernance, c'est tout d'abord un partage clair des intérêts et objectifs de chacun, comme ceux de l'ensemble du collectif. Le projet de gouvernance des données ne déroge pas aux bonnes pratiques de n'importe quel autre projet de gouvernance. Tout au long de ce travail, nous avons exploré les spécificités de la gouvernance des données. À travers de nombreux exemples de terrain, nous avons tâché de montrer leur diversité et la manière dont elles peuvent conduire à gagner en efficacité et à créer de nouvelles valeurs.

Le temps où la donnée était jalousement gardée est désormais révolu. Peu d'acteurs peuvent se prévaloir d'être les uniques propriétaires d'une donnée : la valeur de celle-ci prend tout son sens dans sa circulation et son échange. Le projet de gouvernance des données permet à l'organisation de nouer des alliances stratégiques pour faire levier sur son activité. Matière plastique et vivante, la donnée se prête si bien à cette nouvelle forme d'usage et de partenariat. Et vous, où en est votre organisation ?

ÉVALUATION DE LA GOUVERNANCE DES DONNÉES

Cette partie propose quelques document-outils facilitant l'évaluation et le suivi du projet de gouvernance.

Fiche d'identité de la gouvernance

Nom du projet	
Responsable(s)	Prénom(s) Nom(s), rôle, organisation, équipe, coordonnées
Historique (dates clés)	Date de démarrage Date des rencontres importantes Date de changements notables (objectifs, responsabilités) ...

Périmètre de la gouvernance

Données concernées (orga, service)	La liste des jeux de données s'il y en a peu. Sinon les domaines concernés par les données.
Objectif de la gouvernance	Objectif, moyens, délais.
Principaux attendus	Liste des 3-4 principaux attendus concrets de la gouvernance, si possible avec des délais : livrables et éventuels impacts attendus

Périmètre de la gouvernance

Prénom Nom, email	Organisation et/ou service	Rôle dans la gouvernance

ÉVALUATION DU CATALYSEUR

Pour un projet de gouvernance, il est important d'identifier tous les catalyseurs et d'en hiérarchiser l'importance.

Pour faire ressortir les différences entre les acteurs nous vous conseillons d'utiliser une notation segmentée de un à dix qui accentuera les profils, comme par exemple :

- 2 : catalyseur qui n'agit pas sur la motivation de l'acteur
- 5 : catalyseur qui fait participer l'acteur à la gouvernance
- 7 : catalyseur qui stimule particulièrement la motivation de l'acteur
- 10 : catalyseur qui rend l'acteur moteur dans les actions en lien direct avec ce catalyseur

	Catalyseurs			
	Opportunités financières	Création de valeur pour les acteurs de la gouvernance	Création de valeur pour un éco-système	Contraintes juridiques
Acteur 1				
Acteur 2				
Acteur 3				
Acteur 4				
Acteur 5				
...				
Moyenne pour la gouvernance				
Stratégie partenariale retenue				

MATURITÉ DES POLITIQUES DATA

Ce tableau permet au responsable du projet de gouvernance d'avoir une vision synthétique de l'état de l'art des politiques relatives aux données. Il peut être renseigné pour la gouvernance multi partenariale ou pour chacune des organisations.

Le niveau de maturité peut s'exprimer sur 3 niveaux : faible, moyen, élevé.

Définition des politiques de données...	Niveau de maturité	Nom de la personne responsable
de formation en matière de données : gestion, qualité, usages, etc.		
organisationnelle en matière de données : enjeux, fonctions, métiers, profils, postes, partenariats, charte, etc.		
en matière de création/suppression de données et conditions d'utilisation (lecture, écriture, utilisation)		
juridique de l'entreprise sur les sujets "donnée" (process de validation des textes juridiques, règles globales, etc.)		
de sécurité des données : politique d'accès, politique de protection, etc.		
de gestion technique des données : stockage, exposition, moyens d'accès, etc		
de qualité des données		
de sauvegarde		
des expérimentations		
en matière d'archivage de données		
de communication et de valorisation de la donnée (les animations, la culture des données, les supports et les canaux de communication)		
de (selon les métiers de l'organisation)		
de (selon les métiers de l'organisation)		

CARTOGRAPHIE DES EXPERTISES DES DONNÉES DANS L'ORGANISATION

Le tableau ci-dessous s'appuie sur le cycle de vie des données. Il liste tous les traitements qui sont opérés sur les données. Il permet de repérer qui s'occupe de quoi concernant les données.

Cycle de vie des données	Opérations associées au cycle de vie des données	Quelle organisation / Quel service	Qui
Urbaniser	Construire la cartographie des données et les dictionnaires des données		
	Architecturer et urbaniser le SI (serveurs, tuyaux, protocoles, etc.)		
	Mettre les données en conformité avec les réglementations et standards (Inspire, etc.)		
Créer Adminis- trer	Modéliser : être capable de mettre en œuvre une structure logique destinée à contenir des données. Définir leur cycle de vie de manière à faciliter leur collecte, leur enregistrement ou leur traitement.		
	Structurer les données : définir une méthode de collecte en fonction de la structure de la base de données. Optimiser le processus de saisie. Permettre des modes d'imports de données personnalisés et adaptés aux usages.		
	Gérer : utiliser un système de gestion de base de données, et assurer l'ensemble des opérations d'administration. Caractériser dans le temps les différents statuts de la donnée		
Mettre à jour Contribuer	Recueillir : générer des données de qualité pour mes propres besoins à partir des activités ou pour répondre à la sollicitation d'un tiers.		
	Editer : s'engager et coopérer pour produire des données en utilisant des outils, plus ou moins complexes, d'édition collaborative et ajouter mes contributions dans une base commune et distribuée, administrée par d'autres.		
	Améliorer : contribuer à l'amélioration d'une base de données. Contrôler la qualité des données et détecter des erreurs. Corriger des données, les compléter avec des imports, et les rendre réutilisables.		
Piloter Interpréter	Contextualiser : savoir repérer et vérifier les informations associées aux données, notamment les notices de description (contenu, légende, méthode de collecte...), les conditions de production et de réutilisation, les formats, etc.		
	Analyser : être capable d'explorer et de questionner des données, de qualifier leur nature, de comprendre leur structure, de les trier / filtrer pour dégager des clés de lecture et d'effectuer des traitements simples pour faciliter l'analyse.		
	Evaluer : À partir de l'analyse, apprécier la qualité et la pertinence des données et anticiper leurs manques. Mesurer / écarter leurs biais et identifier les éventuelles erreurs à corriger pour les exploiter.		
Utiliser et échanger Accéder + Trans- former + Valoriser	Comprendre les données : déterminer les données utiles pour satisfaire un besoin. Fixer des critères et un périmètre pour définir des sources pertinentes où obtenir ces données		
	Rechercher : rechercher des informations et des données en ligne, utiliser des fonctions avancées pour affiner les recherches, sélectionner des sources et des producteurs, fouiller dans les portails et les sites spécialisés.		
	Obtenir : mettre en œuvre la méthode d'obtention la plus adaptée : téléchargement, extraction (doc, web, api), réclamation, achat.		
	Nettoyer : modifier, compléter, réorganiser, normaliser et corriger des données pour assurer l'exécution de traitements automatiques ou pour les rendre compatibles avec un logiciel.		

	Traiter : opérer l'ensemble des traitements (fonctions, formules, conversions, géocodage) qui permettent de préparer et de transformer des données en vue de leur analyse ou de leur valorisation.		
	Combiner : réconcilier des données issues de sources hétérogènes, statiques ou dynamiques, de les croiser et de les agréger à partir d'opérations plus ou moins automatisées, pour les utiliser comme un tout unifié.		
	Représenter : en fonction du type des données (statistiques, spatiales, temporelles...) et des informations issues de leur analyse, choisir des outils et produire une forme de représentation adaptée (graphique, carte, timeline...)		
	Calculer : choisir / élaborer et implémenter un modèle statistique, plus ou moins complexe, dans un tableur ou un système de base de données. Comprendre les instructions de calcul et leurs implications sur les résultats.		
	Créer : concevoir et développer une application, web / mobile, en s'appuyant sur des données, proposer un service à valeur ajoutée dont l'architecture est fondée sur l'exploitation créative de ces données		
Publier et diffuser Communiquer	Présenter : sélectionner et traduire des données en informations, les mettre en forme pour les rendre intelligibles, les commenter, les expliquer et faire passer des messages clés issus de leur analyse.		
	Éditorialiser : scénariser et raconter une histoire à partir de données, les représenter sous des formes visuelles adaptées (mêlant rationnel et sensoriel) et les rendre accessibles grâce à une narration compréhensible.		
	Débattre ou décider : instaurer et d'éclairer un débat à partir de données, dialoguer jusqu'à prendre (aider à prendre) une décision, favoriser et orienter le passage à l'action en s'appuyant sur l'analyse de ces données.		
	Répertorier : détailler et décrire précisément la structure d'un jeu de données (valeur, série), lui attribuer des métadonnées (nature, date, production...). Les introduire de façon manuelle ou plus ou moins automatisée.		
	Standardiser : respecter les normes et standards d'échange de données. Publier des données interopérables. Anticiper les conséquences du système dans lequel les données seront intégrées. Privilégier des formats ouverts et lisibles.		
	Diffuser : rendre des données disponibles en téléchargement ou les exposer via une API pour offrir un accès fiable aux développeurs et autoriser le moissonnage par des tiers. Inciter / encourager à réutiliser les données publiées.		
Archiver	Structurer/Administrer les bases de données de son organisation		
	Favoriser l'interopérabilité des données et leur libre circulation		
	Former les publics		
	Collecter et classer les données		
	Supprimer des données (délais légaux, procédures...)		

CAHIER #3

Nouvelles efficacités et création de valeur : les projets de gouvernance des données

<http://infolabs.io/gouv16>



CC License Creative Commons BY 3.0

Ce cahier est une production de la Campagne Infolab de la Fing



avec le soutien des organisations participantes



La Fing a le soutien de ses grands partenaires

