



HAL
open science

La gouvernance des données en contexte universitaire : proposition d'un modèle de maturité

Ugo Verdi, Nathalie Pinède, Guy Melançon

► To cite this version:

Ugo Verdi, Nathalie Pinède, Guy Melançon. La gouvernance des données en contexte universitaire : proposition d'un modèle de maturité. INFORSID 2024, May 2024, Nancy, France. pp.21-36. hal-04585975

HAL Id: hal-04585975

<https://hal.science/hal-04585975>

Submitted on 23 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La gouvernance des données en contexte universitaire : proposition d'un modèle de maturité

Ugo Verdi^{1,2}, Nathalie Pinède¹ et Guy Melançon²

1. Laboratoire MICA, Université Bordeaux-Montaigne
Bâtiment MSH de Bordeaux, 10 Esplanade des Antilles, 33607 Pessac cedex
ugo.verdi@etu.u-bordeaux-montaigne.fr et nathalie.pinede@u-bordeaux-montaigne.fr

2. Laboratoire LaBRI, Université de Bordeaux
Domaine universitaire, 351 cours de la Libération, 33405 Talence
guy.melancon@u-bordeaux.fr et ugo.verdi@u-bordeaux.fr

RÉSUMÉ. Dans le contexte d'une production très rapide des données, la question de leur gouvernance traverse l'ensemble des organisations. Si cette problématique a été principalement étudiée dans un contexte entrepreneurial où une gestion optimale des données sert en priorité des objectifs économiques, elle a moins fait l'objet d'une étude dans un cadre universitaire. Cet article propose ici un modèle de maturité pour permettre aux universités de disposer d'un outil concourant d'une part à identifier et analyser les mécanismes de la gouvernance des données et, d'autre part, à évaluer le niveau d'implication des acteurs et de l'organisation. Bien qu'adapté à un contexte local particulier, à savoir les projets ACT de l'université de Bordeaux, il a pour vocation d'être adaptable à d'autres contextes universitaires.

ABSTRACT. As data production continues to surge, the challenge of data governance becomes increasingly pertinent across all organizations. Although extensively explored within business sectors, primarily for its economic benefits in data management, this topic has garnered less attention within university settings. This study introduces a maturity model specifically designed for universities. It aims to identify and analyze data governance mechanisms and assess stakeholder engagement and organizational involvement. Developed within the context of ACT projects at the University of Bordeaux, this model is designed to be adaptable to other university environments.

Mots-clés : modèle de maturité, gouvernance des données, enseignement supérieur

KEYWORDS: maturity model, data governance, higher education

1. Introduction

Dans le contexte d'une production très rapide des données¹, la question de leur gouvernance traverse l'ensemble des organisations. Si cette problématique a été principalement étudiée dans un contexte entrepreneurial (Jimenez et al., 2019) où une gestion optimale des données sert en priorité des objectifs économiques, elle a

¹ <https://www.statista.com/statistics/871513/worldwide-data-created/>

moins fait l'objet d'études dans un cadre universitaire. Or, l'absence de gouvernance expose indifféremment les organisations à de nombreux risques d'ordre réglementaire, sécuritaire et logistique qui entravent la bonne gestion des données et amènent à questionner leur fiabilité et légitimité (Verdier, 2015 ; Al-Ruithe, Benkhelifa et Hameed, 2016 ; Kremser et Brunauer, 2019) ; ce que Borgman et Brand (2022) ont nommé une « taxe invisible » sur l'efficacité. Le monde universitaire anglophone s'est emparé de cette thématique en inscrivant la gouvernance des données dans leur fonctionnement depuis une dizaine d'années (Jim et Chang, 2018). En France, les initiatives visant la mise en place d'une gouvernance des données sont assez confidentielles². De plus, les modes d'expression de la gouvernance universitaire en France, bien plus décentralisée que celles du monde anglophone (Melançon, Pinède et Verdi, 2024), ajoutent de la difficulté à une situation naturellement compliquée. En effet, la mise en place d'une gouvernance des données est complexe et soulève des problématiques diverses. En amont de la nécessité de les résoudre se pose celle de les identifier. Cet article s'inscrit dans le cadre d'une recherche action, celle du projet GouD³, afin de proposer un modèle de maturité pour permettre aux universités de disposer d'un outil concourant d'une part à identifier et analyser les mécanismes de la gouvernance des données et, d'autre part, à évaluer le niveau d'implication des acteurs et de l'organisation. Bien qu'adapté à un contexte local particulier, à savoir les projets ACT de l'université de Bordeaux, il a pour vocation d'être adaptable à d'autres contextes universitaires.

2. La gouvernance des données en contexte universitaire

Dans nos précédents travaux (Melançon et Pinède, 2023 ; Verdi, 2023a), nous avons défini la gouvernance des données comme l'exercice de l'autorité et du contrôle sur la gestion des données par l'institution d'un système de normes et de procédures (Plotkin, 2013). Celle-ci vise une exploitation optimale des données, notamment grâce à une politique générale d'ouverture des données (Gegenhuber et al., 2023), afin d'aiguiller les décisions de l'organisation (Janssen, 2020). Dans ce cadre, la gouvernance doit s'assurer du respect de la conformité de ses principes (Loshin, 2008), ce qui est souvent englobé dans le vocable anglais de *compliance* (Fu et al., 2011). Ce système de principes et de règles, intégré dans une stratégie globale de long terme (Weber, Otto et Österle, 2009 ; Waller, 2020), induit la constitution d'une culture commune sur les données où ces dernières sont valorisées et perçues comme porteuses de valeur pour les acteurs de l'université. Son émergence est grandement liée à la culture des données (*data literacy*) diffusée au sein de l'université. Cette dernière comprend à la fois des connaissances théoriques et techniques mais intègre également un aspect de formation (Verdi, 2023b). Elle est

² Le seul véritable exemple est celui de l'Université Côte d'Azur (voir AMUE, 2022).

³ Le projet GouD est l'un des projets du programme ACT (*Augmented university for Campus and world Transition*) de l'université de Bordeaux qui mène un travail prospectif sur la gouvernance des données et examine l'état de l'art sur les questions de gouvernance pour cerner les particularités du contexte universitaire français et *in fine* proposer un modèle de gouvernance spécifiques aux établissements ESRI.

essentielle à la gouvernance en permettant « l’enculturation » des acteurs qui, au-delà d’une simple acculturation, favorise le mécanisme d’accomplissement du changement et amène à une acceptation plus naturelle de ses principes (Herskovits, 1952). Culture des données et gouvernance des données sont donc indissociables et forment deux « blocs constitutifs importants du socle de connaissances des professionnels de l’information impliqués dans le soutien des recherches intensives des données »⁴ (Koltay, 2016).

La gouvernance englobe ou est englobée dans un certain nombre de modalités. En premier lieu, la gestion des données (*data management*) : au contraire de la gouvernance qui institue les principes, les règles et les valeurs ainsi que la répartition des rôles des acteurs qui la composent, la gestion des données a pour but la mise en place effective de ceux-ci (Khatri et Brown, 2010 ; Alhassan, Sammon et Daly, 2018 ; Rafal et Girard, 2023). Cette gestion des données se confond souvent dans la littérature avec l’intendance des données (*data stewardship*) qui se présente également comme la facette opérationnelle de la gouvernance (Plotkin, 2013). Dans les deux cas, qu’il s’agisse de la gestion des données ou de l’intendance des données, l’objectif poursuivi est le même : s’assurer de la qualité des données, comprise dans l’activité de la gestion de la qualité des données (*data quality management*) (Wende et Otto, 2007). Celle-ci comprend la conformité de leur complétude, de leur consistance, de leur précision, de leur pertinence, de leur interprétabilité, de leur réutilisabilité ou encore de leur rapidité d’obtention (Pinino, Lee et Yang, 2002 ; Cheong et Chang, 2007 ; Brous, Janssen et Herder, 2016). Cette gestion de la qualité, au-delà de s’assurer de la bonne facture des données, doit prendre en compte la sécurité de ces dernières (*data security*) pour éviter toute fuite ou perte (pouvant porter atteinte à la vie privée des acteurs par exemple), ce qui suppose une accréditation des personnels (Benfeldt, Persson et Madsen, 2020) et la mise en place d’un ensemble de procédures de protection à court et long terme. La difficulté réside ici dans la recherche d’un bon équilibre entre l’accès et le contrôle des données pour ne pas enrayer le fonctionnement de la gouvernance (Rafal et Girard, 2023).

Pour assurer ce fonctionnement, plusieurs typologies d’acteurs sont intégrées à la gouvernance. Nous avons pu lister (1) les administrateurs (*data trustees*) qui sont les responsables de la conformité du respect des règles de la gouvernance pour garantir l’intégrité et l’utilité des données ; (2) les propriétaires (*data owners*) à l’origine de la conception des jeux de données et peuvent être tout aussi bien des personnels internes (ex : enseignant-chercheur) qu’externes à l’université (ex : une institution comme l’INSEE) ; (3) les utilisateurs (*data users*) qui s’emparent des jeux de données pour des usages spécifiques et qui n’en sont pas nécessairement les destinataires originels ; 4) les intendants (*data stewards*) ou managers (*data managers*) qui sont des experts disciplinaires donnant des conseils sur les différents aspects liés à la gestion des données (Teperek et al., 2018). Il est à noter que les personnes occupant les rôles (1) et (2) ne peuvent pas toutes être incluses dans la

⁴ « *Data governance and data literacy are two important building blocks in the knowledge base of information professionals involved in supporting data-intensive research, and both address data quality and research data management* ».

gouvernance pour des raisons pratiques : désigner des représentants pourrait être une solution plus simple pour recueillir leurs contributions.

Dans le monde universitaire français, ces rôles ne sont pas nécessairement constitués comme tels et un acteur peut assurer à lui seul plusieurs de ces fonctions comme dans le cas d'un délégué à la protection des données (DPO). Néanmoins, cette conceptualisation est utile pour délimiter un premier périmètre d'étude. Avec l'aval des instances dirigeantes, ces acteurs peuvent ensuite se constituer en comité, pourvu ou non d'un pouvoir décisionnel. Dans les universités nord-américaines, un même modèle de gouvernance, adapté du DMBOK⁵, est visible : celui-ci est vertical, piloté en premier lieu par le *provost* (que nous pourrions rapprocher du rôle de Directeur Général des Services) qui prend part, voire constitue, un comité de la gouvernance des données (*data governance board* ou *council*) intégrant des acteurs issus de l'ensemble des services de l'université⁶ qui élaborent et supervisent l'application des mécanismes. Au contraire de cette approche descendante qui transforme la gouvernance en un organe de contrôle, nous avons souligné dans Melançon et Pinède (2023) que l'approche « non-intrusive » de Seiner (2014) mettant la gouvernance au service des projets semblait plus adaptée au contexte local français. Dans le cadre de l'université de Bordeaux, cette implémentation débute précisément par un ensemble d'initiatives limitées à l'écosystème des projets ACT que nous allons maintenant présenter.

3. Le contexte bordelais : le projet ACT

Débuté en 2021, le projet *Augmented university for Campus and world Transition* (ACT) est un projet ANR ayant pour mission « de développer, tester, valider et diffuser de nouvelles façons d'aborder les grands problèmes environnementaux, sociaux et de transition économique grâce à la transformation des campus de l'université en un vaste laboratoire vivant »⁷. Il comprend plus de 24 projets, intégrés dans des *living labs*, liés aux questions de transition (écologique, économique et numérique) où interagissent un grand nombre d'acteurs et dans lesquels les productions de données s'insèrent dans des visées différentes. Nous pouvons citer à titre d'exemple les projets « PRISME » manipulant les données de santé d'étudiants, « Datacampus » travaillant sur les données de mobilité du campus, ou encore « Forêt Urbaine » analysant les données issues des réponses de la forêt urbaine au changement climatique. En surplomb se greffe le projet « Datalab », débuté en 2023, dont l'objectif est de développer, à l'échelle des projets ACT, des actions de stratégie data, de gouvernance des données et de développement des

⁵ Le DMBOK est une publication de la DAMA International qui décrit les mécanismes de la gouvernance des données dans un contexte entrepreneurial.

⁶ Nous retrouvons ce modèle dans de nombreuses universités nord-américaines comme celles d'East Carolina, de Madison, du Maine, de San Francisco ou de Villanova, mais également sous une forme similaire dans les universités de Toronto (Canada) et de Londres (Royaume-Uni).

⁷ <https://www.u-bordeaux.fr/universite/notre-strategie/projets-institutionnels/act-pour-un-campus-experimental>

usages de la donnée (Blanchard, 2023). La volonté d'instauration d'une gouvernance des données est néanmoins plus ancienne : elle s'est faite jour officiellement avec la création d'un COPIL « Gouvernance des données et des documents d'activité publics de l'établissement » (GDDP) en 2017 sous l'impulsion de la Direction des archives universitaires, au sein de laquelle une réflexion et des travaux ont menés à la production d'une note interne de contexte sur la gouvernance des données en 2021 et d'un article d'Anne Pletinckx, directrice des Archives universitaires (Pletinckx, 2022), qui proposent tous deux une description des caractéristiques de la gouvernance des données et qui soulignent les prémisses de l'intérêt du sujet au sein de l'université. La gouvernance n'en est actuellement qu'à un stade embryonnaire, entamant depuis peu la poursuite d'initiatives telles que la réalisation de Plans de Gestion des données (PGD), la constitution en décembre 2023 d'un groupe accueillant un certain nombre d'acteurs internes et externes à l'université de Bordeaux contribuant à la définition des principes de la gouvernance et dernièrement avec la création en février 2024 d'un service 3D (Données, décisionnel, *datalab*) ayant en charge l'animation de la gouvernance des données et la diffusion d'une culture des données au sein de l'université. En tant que projet de recherche-action s'appuyant sur une démarche empirique et inductive, GouD utilise l'ensemble des projets ACT précités comme terrain d'expérimentation. Le modèle de maturité, issu de notre lecture de la littérature mais également des observations issues des projets ACT, s'inscrit ici dans une analyse de long terme de l'élaboration de la gouvernance.

4. Le modèle de maturité

Les objectifs poursuivis par notre modèle de maturité sont les suivants : clarifier les caractéristiques générales de la gouvernance des données et pouvoir les analyser à travers le prisme de trois facettes. Pour chaque facette, des composantes seront décrites et des sources permettant de les analyser seront présentées. Le tout pour aider à la mise en place d'une gouvernance des données ou pour en améliorer une déjà existante. Notre modèle de maturité doit donc permettre de poser un constat sur une situation en cours, à différents stades, et anticiper son évolution.

Selon la littérature, un modèle de maturité sert d'outil de diagnostic (Kohlegger, Maier & Thalmann, 2009) ayant été testé dans un contexte spécifique (Wilkinson, 2014) et comprenant deux perspectives potentielles : un cycle de vie ou un potentiel de performance (McBride, 2010). Il peut être construit selon un modèle descriptif, prescriptif ou comparatif (Poeppelbuss et Roeglinger, 2011) et intègre des niveaux de maturité pour un ensemble d'items comme des processus, des rôles ou une organisation par exemple. Ces niveaux représentent une ligne d'évolution anticipée ou désirée (Becker, 2009). Notre modèle doit permettre de déterminer un niveau de maturité de la gouvernance des données mise en place au sein des projets ACT de l'université de Bordeaux. A ce titre, il doit intégrer dans son analyse l'ensemble des acteurs ayant une interaction avec les données employées au sein de l'université : les acteurs des projets ACT mais également ceux des directions de services (ex : la direction des archives universitaires), voire les partenaires extérieurs (ex : la

Métropole de Bordeaux). Les concepts précités (ex : *data management*, *data quality*, *data literacy*, etc.) sont ici ré-agencés pour proposer une grille de lecture adaptée à un contexte français.

La maturité de la gouvernance des données n'est que peu définie dans la littérature, à l'exception de Marchildon et al. (2018) et de Gupta et Cannon (2020). Selon ces auteurs, celle-ci renvoie au niveau de conception et de déploiement de règles et de principes visant à optimiser la gestion des données sur leur cycle de vie, à savoir leur détection, leur production, leur collecte, leur usage, leur partage, leur stockage, leur archivage et leur destruction. Pour être mesurable, elle nécessite l'analyse de deux critères : l'existence ou non d'une variable (ex : les moyens humains) et la capacité de l'université à pouvoir identifier ces variables (ex : les données employées dans un laboratoire). Comme tout modèle de maturité, le nôtre intègre des niveaux qui explicitent la formalisation, la généralisation et l'adhésion aux principes de la gouvernance. Dans la littérature, ils sont généralement au nombre de cinq et leurs intitulés diffèrent selon les sources (Soares, 2013 ; Seiner, 2014 ; Merkus, 2015 ; Lee et al., 2019). Carretero et al. (2016) citent un niveau 0 pour signifier une absence totale d'initiatives. Nous avons fait le choix de six niveaux de maturité : un niveau 0 (inexistante), un niveau 1 (en émergence), un niveau 2 (en développement), un niveau 3 (déploiement local), un niveau 4 (déploiement global) et un niveau 5 (mature). Ces niveaux se déclinent selon trois facettes : les données, les acteurs et l'organisation. Leur sélection s'est faite à partir de la littérature (scientifique et professionnelle) décrivant les caractéristiques de la gouvernance des données mais également à partir d'autres facettes mises en avant par les auteurs (comme les personnes, les normes ou encore les technologies) que nous avons remaniées ainsi que des difficultés rencontrées dans l'élaboration de la gouvernance. Ces deux derniers points, bien que n'étant pas originellement intégrés dans une logique d'analyse de maturité, ont toutefois permis d'affiner notre proposition.

Plusieurs auteurs et rapports ont contribué à lister les freins potentiels et les défis à relever pour pouvoir construire et asseoir une gouvernance des données (DAMA International, 2009 ; Benfeldt, 2017 ; Mahanti, 2021 ; Okoro, 2021). Y sont cités les problèmes (1) de compréhension commune sur ce que recouvre la gouvernance des données, (2) de logistique vis-à-vis de la mise en place et de la conformité de son application, (3) de conflits entre acteurs et/ou services dus à des questions de partage de pouvoir, de périmètres d'action ou encore de visions concurrentes sur les données et leur traitement, (4) de refus du changement découlant d'un rejet des règles et principes de la gouvernance, (5) d'absence de compétences et de moyens (matériels, financiers ou humains) pour traiter les nécessités de la gouvernance. Au point (5), l'absence est due soit à une non-spécialisation des acteurs, soit à une incapacité de les définir et *in extenso* de recruter les personnes compétentes, soit à un désinvestissement de cette problématique par les personnes ayant un pouvoir décisionnel, notamment celles en charge du budget de l'établissement. Ces points sont transversaux et doivent être déduits de l'analyse de chaque composante de notre modèle de maturité, en particulier la non-adhésion aux principes de la gouvernance.

La première facette traite des données. On lui dénombre trois sous-facettes : (1) le capital des données, c'est-à-dire l'ensemble des données existantes et les moyens disponibles pour les recenser et les gérer, (2) les règles et principes encadrant cette gestion et (3) les formations proposées pour établir une culture commune des données.

Tableau 2: la facette des données

Sous-facettes	Composantes	Description des composantes
Le capital	Les données recensées	L'ensemble des données utilisées dans différents projets ou contextes. Les caractéristiques d'un ensemble de données doivent être documentées de manière exhaustive en précisant leur formats, leurs métadonnées, etc.
	Les outils de mesure	Les livrables permettant de lister et surveiller les flux de données (ex : un plan de gestion de données, un <i>dashboard</i> ou un catalogue de données).
	Les moyens	Les caractéristiques des moyens humains (les tâches et les rôles), structurels (ex : les serveurs, les logiciels, les locaux dédiés) et financiers.
Les règles et les principes	Le cycle de vie des données	Les règles et principes encadrant l'ensemble du cycle de vie des données, à savoir l'identification, la collecte, la production, le stockage, l'archivage et la destruction des données (ex : intégrés dans un arbre décisionnel). Si rendus obligatoires, les logiciels et technologies choisis doivent ici être précisés et décrits. Du fait du grand nombre de tâches que cette composante implique, elles peuvent être traitées séparément (dans le cadre d'un audit par exemple).
	La qualité	L'ensemble des règles et principes employés pour s'assurer de la pertinence, de la complétude, de la conformité, de l'intégrité, de la fraîcheur et de la cohérence des données (ex : principes <i>FAIR</i> , norme des métadonnées, etc.).
	La sécurité	L'ensemble des protections matérielles et cyber ainsi que la gestion des accès aux données.
	La documentation	Les livrables listant l'ensemble des règles et des principes de gestion des données.
Les formations	Les formats	Le déroulement des formations (ex : la temporalité, les activités proposées, etc.) et les sujets qu'elles abordent (ex : la science ouverte).
	Les publics visés	Les types de publics (ex: les étudiants, les directions universitaires, etc.) et les raisons de ce ciblage (ex : amélioration de compétences, développement de nouvelles connaissances, etc.).
	Les formateurs	Le profil des formateurs (ex : est-ce un acteur interne / externe, quelle est la nature de son poste, etc.) et la raison de leur recrutement (ex : compétences uniquement détenues par ces derniers, partenariat historique avec l'université, etc.).

La seconde facette traite des acteurs. Son objectif est de souligner le rôle humain en prenant en considération l'autonomie des acteurs et leur aptitude à accepter ou non le changement. Devront ici être soulevés leurs rôles respectifs, leurs méthodes de travail, leurs modes de collaboration ainsi que le degré de leur culture des données. Il s'agit donc ici d'une enquête portant à la fois sur les pratiques effectives et sur les imaginaires des acteurs.

Tableau 3: la facette des acteurs

Sous-facettes	Composantes	Description des composantes
Les typologies	Les administrateurs	Les acteurs en charge du respect de l'application des principes et règles de traitement des données.
	Les producteurs	Les acteurs en charge de la production des données.
	Les utilisateurs	Les acteurs employant les données mises à disposition par l'université.
	Les intendants ou managers	Les acteurs conseillant et supervisant la gestion des données.
Les usages	Les contextes et les objectifs	La description des contextes de production et de traitement des données ainsi que les objectifs que servent les données.
	Le traitement	La description du traitement des données sur l'ensemble de leur cycle de vie (ex : quel acteur traite quelle donnée, selon quelle temporalité, quels matériels, dans quel objectif, etc.).
	La conformité	La capacité de la gouvernance d'intégrer les différents cadres réglementaires (ex : <i>RGPD</i>), de pouvoir dialoguer avec d'autres instances ayant également développé des règles (ex : comité éthique, comité de santé, etc.) et de s'assurer du respect des règles établies.
Les collaborations	Les modalités de collaboration	La manière dont les acteurs collaborent entre eux (ex : leurs méthodologies, leurs connaissances et représentations respectives).
La culture des données	Les niveaux de formation	Les connaissances et compétences des acteurs sur les données et leur écosystème dont fait partie la gouvernance des données.
	Les représentations	Les visions et imaginaires des acteurs sur les données et leur écosystème dont fait partie la gouvernance.

La troisième facette traite de l'organisation et se réfère à l'élaboration et la supervision de la gouvernance. Elle liste l'ensemble des éléments à prendre en considération pour garantir sa pérennité. Elle contient (1) la stratégie qui vise à établir et à faire connaître la vision de la gouvernance défendue par l'université

auprès des acteurs qui la composent, (2) la forme qui la régit et (3) les collaborations entretenues entre acteurs internes et externes à l'université.

Tableau 4: la facette de l'organisation

Sous-facettes	Composantes	Description des composantes
La stratégie	La vision stratégique	Les valeurs et les objectifs globaux de la gouvernance ainsi que la manière dont elle est reçue et interprétée par les acteurs de l'université.
	La communication	Les caractéristiques des stratégies de communication (ex : les acteurs qui en ont la charge, la temporalité des actions de communication, les sujets abordés, etc.) et leur réception par les acteurs (ex : adhésion, rejet, etc.).
La forme	La coordination de la gouvernance	Précise la manière dont la gouvernance est coordonnée (centralisation, décentralisation ou subsidiarité) et les raisons ayant conduit à ce choix.
	L'attribution des rôles	Désigne les mécanismes d'attribution des rôles aux acteurs pour traiter un point spécifique de la gouvernance.
Les collaborations	Les collaborations internes	Les acteurs internes à l'université (ex : directions de services) et les modalités des collaborations (ex : rôle consultatif permanent, temporaire, etc.)
	Les collaborations externes	Les acteurs externes à l'université et les modalités des collaborations.

Au même titre que les mécanismes de la gouvernance décrits par Abraham et al. (2019), notre modèle se veut interactionniste : chaque facette a une influence sur les autres et aucune n'est prévalente. Toutes doivent être développées pour que la gouvernance des données atteigne un niveau de maturité acceptable, déterminé selon les besoins et objectifs de l'université.

L'ensemble des points précédents doit permettre ainsi de déterminer un niveau de maturité pour chacune des facettes à un instant t. En effet, de nombreux changements internes (budgétaires, matériels, logistiques et humains) ont lieu dans le temps, en particulier le *turnover* fréquent des personnels, auxquels s'ajoutent d'autres changements externes comme l'évolution des réglementations. Le niveau de maturité mesuré n'est donc pas définitif et devra être à nouveau analysé.

Dans notre tableau récapitulatif des niveaux par facettes, les cases ne doivent pas se lire de manière indépendante, les composantes d'une facette influant directement celles d'une autre facette. A titre d'exemple, sans actions de communication, les représentations des acteurs sur la gouvernance n'évolueront pas.

Ainsi, dans cette conception, les niveaux ne sont pas homogènes : une facette pourra par exemple être à un niveau 2 tandis qu'une autre sera à un niveau 0. L'évaluation globale du niveau de maturité est appréciative, dépendante des contextes spécifiques des organisations.

Tableau 5: les niveaux de maturité par facette

Niveaux / facettes	Les données	Les acteurs	L'organisation
Niveau 0 Inexistante	Les données ne sont pas identifiées, ni traitées selon des principes et des règles unifiés. Aucune formation aux données n'est proposée à ce niveau.	Les acteurs ne sont pas identifiés. Leur culture des données est trop peu développée pour qu'ils puissent avoir une vision claire de ce que recouvre la gouvernance.	La gouvernance n'est ni définie, ni déployée.
Niveau 1 En émergence	Une volonté d'identifier et de gérer les données selon des principes et des règles unifiés émerge. Des formations aux données ne sont pas encore proposées aux acteurs.	De rares acteurs ainsi que les modalités de leur collaboration et de leur usage des données sont identifiés. Ils ne comprennent pas complètement ce que recouvre la gouvernance.	Le besoin d'une gouvernance émerge. Aucune stratégie de communication n'est encore élaborée.
Niveau 2 En développement	Quelques données commencent à être identifiées. L'acculturation aux données passe par des formations externes selon les opportunités rencontrées par les acteurs.	Quelques acteurs ainsi que les modalités de leur collaboration et de leur usage des données sont identifiés. Ils commencent à comprendre ce que recouvre la gouvernance.	La gouvernance est cours d'élaboration. Quelques actions de communication sont potentiellement entreprises pour en discuter.
Niveau 3 Déploiement local	Une partie seulement des données est identifiée et est gérée selon les principes et les règles de la gouvernance. Des formations aux données sont proposées ponctuellement par la gouvernance pour soutenir les acteurs.	Une minorité d'acteurs ainsi que les modalités de leur collaboration et de leur usage des données sont identifiés. Leur culture des données développée leur permet de comprendre ce que recouvre la gouvernance.	La gouvernance est déployée à un niveau local, soutenue par une première stratégie de communication qui met à disposition une documentation détaillant ses principes et ses visées.

Niveau 4 Déploiement global	La majorité des données sont identifiées. Celles-ci sont gérées selon les principes et des règles de la gouvernance. Des formations sont mises en place de manière régulière par la gouvernance pour soutenir les acteurs.	La majorité des acteurs ainsi que les modalités de leur collaboration et de leur usage des données sont identifiés. Leur culture des données développée leur permet de comprendre ce que recouvre la gouvernance.	La gouvernance est déployée à un niveau global et une stratégie de communication est désormais entreprise de manière constante. Elle rencontre néanmoins encore des difficultés d'application.
Niveau 5 Mature	L'ensemble des données sont identifiées et gérées selon les principes et règles de la gouvernance. Des formations sur les données sont proposées par la gouvernance de manière constante pour répondre aux besoins des acteurs.	La totalité des acteurs ainsi que les modalités de leur collaboration et de leur usage des données sont identifiés. Ils comprennent et adhèrent aux principes et règles de la gouvernance.	La gouvernance est déployée à l'ensemble de l'université et sa comitologie est étendue aux partenaires de son organisation. Elle est robuste et capable de s'adapter à de nouveaux contextes.

Comme précisé en introduction, l'une des plus grandes difficultés réside dans l'identification d'éléments constitutifs d'une gouvernance, en particulier lorsque ce terme n'est pas employé : soit parce qu'il ne recouvre pas pour les acteurs une volonté de stratégie globale, soit parce qu'il est tout simplement délaissé pour un autre plus représentatif des activités entreprises comme la science ouverte. Or, des initiatives autour des données peuvent constituer le terreau de la gouvernance et doivent en ce sens faire l'objet d'analyses. C'est le cas des ateliers de la donnée qui se positionnent « comme le point d'entrée en proximité des équipes de recherche sur toute nature de besoin relatif à la donnée »⁸ et dont la constitution en réseau facilite les collaborations et l'instauration d'une vision commune sur les données. En nous basant sur la littérature et nos retours de terrain, nous proposons ici trois grandes sources d'analyses pour abonder notre modèle de maturité pour la gouvernance des données.

D'une part, les sources documentaires qui listent les principes et les règles sur le traitement des données. Celles-ci qui s'incarnent dans de nombreux formats. L'on peut ici citer les chartes, référentiels et règlements intérieurs, les arbres décisionnels, les comptes rendus de réunions, les contenus web ainsi que toute documentation

⁸ https://www.ouvrirlascience.fr/wp-content/uploads/2021/10/2021.10.11_AMI_Ateliers-de-la-donne%CC%81e.pdf

annexe qui permettent de déceler les règles en application au sein de l'université et les visions sous-jacentes dans la gestion des données. Assister à leur élaboration apporte une possibilité supplémentaire d'analyser l'évolution des pensées et les modes d'interaction entre acteurs.

Nous recommandons en particulier l'étude des Plans de Gestion des Données (PGD). De par leur description des jeux de données, ils offrent un panorama complet sur le capital de l'université. De plus, ils permettent également d'identifier les acteurs des projets et leurs rôles respectifs. Bien que centrés autour des données de la recherche, ils peuvent toutefois être déclinés pour d'autres types de données et couvrir un plus large ensemble de ressources. Leur élaboration rend possible le questionnement des acteurs pour obtenir des détails plus précis sur leurs méthodes de travail et sur leurs interactions. Bien qu'étant une source documentaire, la singularité de leurs caractéristiques en font à notre sens un élément d'analyse dissocié des autres sources documentaires.

D'autre part, les sources d'enquêtes des acteurs qui peuvent être réalisées soit avec des questionnaires, soit avec des entretiens. A titre d'exemple, Marchildon et al. (2018) ont proposé un questionnaire à choix multiples constitué de 11 items composés de 72 questions permettant d'analyser chaque aspect d'une gouvernance existante. Si les questionnaires permettent des analyses quantitatives immédiates, les entretiens offrent eux un approfondissement des points peu développés dans les documentations précitées. C'est particulièrement le cas de la culture des acteurs qui ne peut émerger que par ce type d'enquête. Cette culture, par-delà les compétences, connaissances et représentations, induit des pratiques spécifiques, notamment dans le cadre d'une discipline ou d'un service instauré depuis longtemps.

Décerner ces pratiques et contraintes est ici primordial pour adapter la gouvernance aux pratiques locales et éviter qu'elle ne soit perçue comme un doublon incohérent ou comme une charge de travail supplémentaire par les acteurs. Savoir qui interroger relève d'une autre difficulté, en particulier sans historique des précédentes actions réalisées. Dans notre cas, les acteurs liés aux projets ACT d'une part et les acteurs liés à la constitution de la gouvernance de l'université de Bordeaux d'autre part ont pu être identifiés en amont de ce travail. Mais nous gardons à l'esprit que d'autres acteurs ayant un intérêt à la gouvernance et ne s'étant pas encore manifestés peuvent avoir mis en place un certain nombre d'initiatives en ce sens, nécessitant alors le déploiement d'une nouvelle vague d'entretiens. Ces trois sources peuvent être sollicitées pour l'ensemble des facettes mais certaines donneront des réponses plus précises que d'autres, c'est pourquoi nous recommandons ci-dessous les plus appropriées.

Tableau 6: les sources recommandées par facettes

Facettes	Sous-facettes	Sources documentaires	Plan de gestion des données	Sources d'enquêtes
Les données	Le capital	limitée	✓	optionnelle
	Les règles et les principes	✓	✓	optionnelle
	Les formations	✓	hors périmètre	✓
Les acteurs	Les typologies	✓	✓	✓
	Les usages	✓	✓	optionnelle
	Les modalités de collaboration	✓	✓	✓
	La culture des données	limitée	peu adaptée	✓
L'organisation	La stratégie	✓	hors périmètre	✓
	La forme	✓	hors périmètre	✓
	Les collaborations	✓	hors périmètre	✓

5. Conclusion

L'édification et la pérennisation d'une gouvernance des données est un processus long et complexe, en particulier en contexte universitaire. A l'heure où la gestion des données revêt une importance cruciale, il devient nécessaire d'en maîtriser les tenants et aboutissants. La synthèse de la littérature a eu pour objet de proposer une caractérisation de la gouvernance et appelle ici une validation. Le modèle de maturité présenté est actuellement testé dans le contexte d'ACT. La lecture de la documentation, dont les PGD, et la réalisation d'enquêtes entreprises tout au long de l'année 2024 devront permettre de décrire les caractéristiques de la gouvernance des données de l'université de Bordeaux et d'en mesurer le niveau de maturité. Elles pourront amener à une évolution des facettes et de leurs composantes afin de mieux cerner les particularités de la gouvernance des données en contexte universitaire. Ceci fera l'objet de futures publications.

Ce travail a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du programme d'Investissements d'avenir portant la référence ANR-20-IDES-0001

Bibliographie

- Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance : A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, 49, 424438. <https://doi.org/10.1016/j.ijinfomgt.2019.07.008>
- Alhassan, I., Sammon, D., & Daly, M. (2018). Data governance activities : A comparison between scientific and practice-oriented literature. *Journal of Enterprise Information Management*, 31(2), 300316. <https://doi.org/10.1108/JEIM-01-2017-0007>
- Al-Ruithe, M., Benkhelifa, E., & Hameed, K. (2016). A Conceptual Framework for Designing Data Governance for Cloud Computing. *Procedia Computer Science*, 94, 160167. <https://doi.org/10.1016/j.procs.2016.08.025>
- Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009). Developing Maturity Models for IT Management. *Business & Information Systems Engineering*, 1(3), 213222. <https://doi.org/10.1007/s12599-009-0044-5>
- Benfeldt, O. (2017). A Comprehensive Review of Data Governance Literature. *Selected Papers of the IRIS, Issue Nr 8 (2017)*, 3, 120133. <https://aisel.aisnet.org/iris2017/3>
- Benfeldt, O., Persson, J. S., & Madsen, S. (2020). Data Governance as a Collective Action Problem. *Information Systems Frontiers*, 22(2), 299313. <https://doi.org/10.1007/s10796-019-09923-z>
- Blanchard, A. (2023). Enjeux de la donnée universitaire, de sa collecte à son exploitation. *Les data de l'ESR : l'enjeu de la maîtrise des données. Risques et outils pour l'audit et le pilotage*. <https://hal.science/hal-04289500>
- Borgman, C. L., & Brand, A. (2022). Data blind : Universities lag in capturing and exploiting data. *Science*, 378(6626), 12781281. <https://doi.org/10.1126/science.add2734>
- Brous, P., Janssen, M., & Herder, P. (2016). Coordinating Data-Driven Decision-Making in Public Asset Management Organizations : A Quasi-Experiment for Assessing the Impact of Data Governance on Asset Management Decision Making. *Social Media : The Good, the Bad, and the Ugly*. Springer International Publishing, p. 573583
- Carretero, A., Freitas, A., Cruz-Correia, R., & Piattini, M. (2016, janvier 1). *A case study on assessing the organizational maturity of data management, data quality management and data governance by means of MAMD*.
- Cheong, L. K., & Chang, V. (2007). *The Need for Data Governance : A Case Study*.
- DAMA International. (2009). *The DAMA Guide to the Data Management Body of Knowledge —DAMA-DMBOK*. Technics Publications, LLC.
- Fu, X., Wojak, A., Neagu, D., Ridley, M., & Travis, K. (2011). Data governance in predictive toxicology : A review. *Journal of Cheminformatics*, 3(1), 24. <https://doi.org/10.1186/1758-2946-3-24>
- Gegenhuber, T., Mair, J., Lührsen, R., & Thäter, L. (2023). Orchestrating distributed data governance in open social innovation. *Information and Organization*, 33(1), 100453. <https://doi.org/10.1016/j.infoandorg.2023.100453>

- Gupta, U., Cannon, S. (2020). Data Governance Maturity Models. In *A Practitioner's Guide to Data Governance* (p. 143165). Emerald Publishing Limited. <https://doi.org/10.1108/978-1-78973-567-320201007>
- Herskovits, M. J. (1952). *Les Bases de l'anthropologie culturelle*. <https://unesdoc.unesco.org/ark:/48223/pf0000244386>
- Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance : Organizing data for trustworthy Artificial Intelligence. *Government Information Quarterly*, 37(3), 101493. <https://doi.org/10.1016/j.giq.2020.101493>
- Jim, C. K., & Chang, H. (2018). The current state of data governance in higher education. *Proceedings of the Association for Information Science and Technology*, 55(1), 198-206. <https://doi.org/10.1002/pra2.2018.14505501022>
- Jimenez, L. M., Polo, J. A., & Duarte, N. A. (2019). Overview of Data Governance in Business Contexts. *IOP Conference Series: Materials Science and Engineering*, 519(1), 012023. <https://doi.org/10.1088/1757-899X/519/1/012023>
- Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, 53(1), 148152. <https://doi.org/10.1145/1629175.1629210>
- Kohlegger, M., Maier, R., & Thalmann, S. (2009). *Understanding Maturity Models. Results of a Structured Content Analysis*. I-KNOW '09 and I-SEMANTICS '09. https://www.researchgate.net/publication/215312013_Understanding_Maturity_Models_Results_of_a_Structured_Content_Analysis
- Koltay, T. (2016). Data governance, data literacy and the management of data quality. *IFLA Journal*, 42(4), 303312. <https://doi.org/10.1177/0340035216672238>
- Kremser, W., & Brunauer, R. (2019). Do we have a Data Culture? In P. Haber, T. Lampoltshammer, & M. Mayr (Éds.), *Data Science – Analytics and Applications* (p. 8387). Springer Fachmedien. https://doi.org/10.1007/978-3-658-27495-5_11
- Lee, Y., Park, W., Shin, D., & Won, Y. (2019). A Study on Data Governance Maturity Model and Total Process for the Personal Data Use and Protection. *Journal of the Korea Institute of Information Security & Cryptology*, 29(5), 11171132. <https://doi.org/10.13089/JKIISC.2019.29.5.1117>
- Loshin, D. (2008). *Master Data Management*. Morgan Kaufmann.
- Mahanti, R. (2021). *Data Governance Success : Growing and Sustaining Data Governance*. Springer. <https://doi.org/10.1007/978-981-16-5086-4>
- Marchildon, P., Bourdeau, S., Hadaya, P., & Labissière, A. (2018). Data governance maturity assessment tool : A design science approach. *Projectics / Proyética / Projectique*, 20(2), 155193. <https://doi.org/10.3917/proj.020.0155>
- McBride, T. (2010). Organisational theory perspective on process capability measurement scales. *Journal of Software Maintenance and Evolution: Research and Practice*, 22(4), 243254. <https://doi.org/10.1002/spip.440>
- Melançon, G., Pinède, N., et Verdi, U. (2024). Redefining Data Governance: Insights from the French University System. *Actes du colloque ICEIS 2024*, Angers.

- Melançon, G., & Pinède, N. (2023). Gouvernance des données et intelligibilité : Une approche méthodologique en contexte universitaire. *Communication & Organisation*, 64(2), 6781. <https://doi.org/10.4000/communicationorganisation/12664>
- Merkus, J. (2015). *Data Governance Maturity Model*. <https://doi.org/10.13140/RG.2.2.19274.16321>
- Okoro, R. (2021). Proposed Data Governance Framework for Small and Medium Scale Enterprises (SMEs). *All Graduate Theses, Dissertations, and Other Capstone Projects*. <https://cornerstone.lib.mnsu.edu/etds/1126>
- Pinino, L. L., Lee, W. W., & Yang, R. Y. (2002). Data quality assessment. *Communications of the ACM*, 45(4).
- Pletinckx, A. (2022). De la gouvernance des données publiques. *Archivistes ! La lettre de l'Association des archivistes français*, 140, 30-31.
- Plotkin, D. (2013). *Data Stewardship : An Actionable Guide to Effective Data Management and Data Governance*. Newnes.
- Poepplbuss, J., & Roeglinger, M. (2011). *What makes a useful maturity model? A framework of general design principles for maturity models and its demonstration in business process management*. 19th European Conference on Information Systems, ECIS 2011.
- Rafal, O., & Girard, D. (2023). *Maîtriser la #data : Un enjeu majeur de 2023*. WENVISION. <https://www.wenvision.com/maitriser-la-data-un-enjeu-majeur-de-2023/>
- Seiner, R. S. (2014). *Non-Invasive Data Governance : The Path of Least Resistance and Greatest Success*. Technics Publications.
- Soares, F. S. F., & De Lemos Meira, S. R. (2013). *An Agile Maturity Model for Software Development Organizations*. 4.
- Teperek, M., Cruz, M. J., Verbakel, E., Böhmer, J., & Dunning, A. (2018). Data Stewardship Addressing Disciplinary Data Management Needs. *International Journal of Digital Curation*, 13(1), Article 1. <https://doi.org/10.2218/ijdc.v13i1.604>
- Verdi, U. (2023a). L'éthique des données dans les chartes éthiques des collectivités territoriales. *Communication et organisation*, 64.
- Verdi, U. (2023b). Quelle(s) réponse(s) à l'enjeu d'acculturation aux données ? Un état de l'art des caractéristiques de la data literacy. *Revue française des sciences de l'information et de la communication*, 26.
- Verdier, H. (2015). *Administrateur général des données—Rapport au Premier ministre sur la gouvernance de la donnée 2015 : Les données au service de la transformation publique* (p. 52). Secrétariat général pour la modernisation de l'action publique.
- Waller, D. (2020). 10 Steps to Creating a Data-Driven Culture. *Harvard Business Review*.
- Weber, K., Otto, B., & Österle, H. (2009). One Size Does Not Fit All—A Contingency Approach to Data Governance. *Journal of Data and Information Quality*, 1(1), 4:1-4:27.
- Wende, K., & Otto, B. (2007). *A contingency approach to data governance*.
- Wilkinson, N. (2014). *A framework for organisational governance maturity : An internal audit perspective*.