



Partage des données, les enjeux des projets réglementaires

Synthèse de conférence

Université Paris Dauphine-PSL, 29 septembre 2022



Dauphine | PSL 
CHAIRE GOUVERNANCE
ET RÉGULATION

Conférence organisée par la Chaire Gouvernance et Régulation et le Conseil Général de l'Économie

Le 29 septembre 2022



CONSEIL GÉNÉRAL DE L'ÉCONOMIE
DE L'INDUSTRIE, DE L'ÉNERGIE ET DES TECHNOLOGIES



Synthèse n°66
Université Paris Dauphine-PSL

Partage des données, les enjeux des projets réglementaires

Intervenants

Denis Berthault | Président, Groupement français de l'industrie de l'information (GFII)

Éric Brousseau | Directeur scientifique, Chaire Gouvernance et Régulation & Club des Régulateurs, Université Paris Dauphine-PSL

Jordan Cartier | Secrétaire général, Autorité de régulation des transports

Pierre Chastanet | Directeur de l'unité Cloud & Logiciels, DG Connect - Commission européenne

Julien Chiaroni | Directeur du Grand Défi sur « l'IA de confiance pour l'industrie », Secrétariat Général à l'Investissement (SGPI)

Renaud Font | Directeur des opérations, Agdatahub

Claude Kirchner | Directeur de recherche émérite, INRIA (Institut national de recherche en science et technologies du numérique)

Laurent Lafaye | Co-CEO, Dawex

Luc Rousseau | Vice-président du Conseil général de l'économie, de l'industrie, de l'énergie et des technologies au Ministère de l'économie, des finances et de la souveraineté industrielle et numérique

Anne-Sophie Taillandier | Directrice de TerraLab, Institut Mines-Télécom

& Alban Schmutz | Co-fondateur de CISPE (Cloud Infrastructure Services Providers in Europe)

Jean-Yves Ollier | Conseiller d'État, Coordonnateur du n°18 de la revue Enjeux numériques

Clément Tonon | Auditeur au Conseil d'État

Robert Viseur | Chargé de cours et chef de service TIC (Technologies de l'Information et de la Communication), Université de Mons, Belgique

Sommaire

Introduction Éric Brousseau Directeur scientifique, Chaire Gouvernance et Régulation & Club des Régulateurs, Université Paris Dauphine-PSL.....	8
Panorama des politiques et des règlements en cours de développement au sein de l'Union européenne Keynote Pierre Chastanet Directeur de l'unité Cloud & Logiciels, DG Connect - Commission européenne.....	9
Table ronde 1 : Les enjeux du partage des données Modération : Éric Brousseau Directeur scientifique, Chaire Gouvernance et Régulation & Club des Régulateurs, Université Paris Dauphine-PSL.....	12
Les avancées en matière de partage des données de mobilité dans les transports Jordan Cartier Secrétaire général, Autorité de régulation des transports.....	13
Les initiatives françaises et européennes en matière de gouvernance des données de santé Clément Tonon Auditeur au Conseil d'État.....	15
Les projets dans le secteur agricole et les enjeux industriels dans le contrôle des chaînes de production agricole et agroalimentaire, ainsi que sur les marchés des matières premières agricoles Renaud Font Directeur des opérations, Agdatahub.....	17
L'approche trans-sectorielle de l'analyse des enjeux à partir du rôle des données dans l'entraînement des algorithmes Julien Chiaroni Directeur du Grand Défi sur « l'IA de confiance pour l'industrie », Secrétariat Général à l'Investissement (SGPI).....	19
Avantages et inconvénients des modèles alternatifs de rémunération du partage Denis Berthault Président, Groupement français de l'industrie de l'information (GFII).....	21
Conclusion des intervenants.....	23

Table ronde 2 : Comment organiser le partage des données ?

Modération : Jean-Yves Ollier | Conseiller d'État, Coordonnateur du n°18 de la revue

Enjeux numériques.....24

Les principales fonctions des opérateurs de plateformes et fournisseurs de services de partage de données. Quels sont les changements portés par le DGA ?

Laurent Lafaye | Co-CEO, Dawex.....25

Le développement de la gouvernance de l'infrastructure de cloud et les éléments critiques pour la promotion du partage des données

Anne-Sophie Taillandier | Directrice de TerraLab, Institut Mines-Télécom & Alban Schmutz | Co-fondateur de CISPE (Cloud Infrastructure Services Providers in Europe).....27

Les leçons à tirer du modèle de gouvernance du logiciel libre pour l'organisation du partage de données

Robert Viseur | Chargé de cours et chef de service TIC (Technologies de l'Information et de la Communication), Université de Mons, Belgique.....29

L'organisation du partage entre les différentes parties prenantes et les enjeux éthiques

Claude Kirchner | Directeur de recherche émérite, INRIA (Institut national de recherche en science et technologies du numérique).....31

Échanges.....32

Intervention de clôture

Luc Rousseau | Vice-président du Conseil général de l'économie, de l'industrie, de l'énergie et des technologies au Ministère de l'économie, des finances et de la souveraineté industrielle et numérique.....36

Introduction

Éric Brousseau | Directeur scientifique, Chaire Gouvernance et Régulation & Club des Régulateurs, Université Paris Dauphine-PSL

La période est marquée par un mouvement de régulation du numérique de part et d'autre de l'Atlantique, notamment en Europe où de nombreux textes font l'objet de discussions en cours de finalisation à Bruxelles autour des plateformes structurantes (Digital Market Act), des contenus (réseaux sociaux et médias, Digital Service Act) et des données, qu'il s'agisse dans ce dernier cas de l'accès (Data Act) ou de la gouvernance (Data Governance Act) de ces données.

La gouvernance des données couvre différents domaines : les règles pour définir les régimes d'accès et les contrôles d'usage ; les droits des parties prenantes et la négociabilité de ces droits ; le partage de la valeur qui peut être créée à partir des données ; les questions de responsabilité, de sécurisation et d'obligations des différentes parties prenantes. Elle se traduit, en pratique, par une articulation d'opérations techniques et de règles juridiques, d'interventions publiques et privées, de la définition — négociée et imposée — de différents régimes, de normes locales et transnationales.

Les régimes de gouvernance des données sont, pour la plupart, émergents et sont marqués par d'importantes innovations. Les enjeux qu'ils soulèvent autour de la création et de la répartition de valeur sont considérables, qu'ils soient économiques, de puissance ou de souveraineté. C'est la raison pour laquelle on observe une forte concurrence entre acteurs et dans la création de modèles et dans leur mise en œuvre.

L'analyse de ces phénomènes se trouve encore complexifiée par la diversité des données, des usages, des enjeux et des modes de valorisation en fonction des branches d'activité concernées et des périmètres considérés (national, européen, transatlantique, mondial).

Panorama des politiques et des règlements en cours de développement au sein de l'Union européenne

Keynote

Pierre Chastanet | Directeur de l'unité Cloud & Logiciels, DG Connect
- Commission européenne

Au début de l'actuel mandat de la Commission européenne, l'attention était focalisée sur la nécessité de réglementer l'intelligence artificielle (IA) au niveau européen. Insufflé par la vision du commissaire Breton, cela impliquait de maîtriser les données en Europe qui alimentent en amont les processus de machine learning et de stimuler le partage de données entre les différents acteurs des chaînes de valeur industrielle, pour permettre le bon fonctionnement de l'IA et le développement de l'économie de la donnée autour des écosystèmes industriels européens. C'est ainsi qu'est né l'un des tout premiers documents publiés par la Commission européenne au cours de ce mandat et qui guide encore son action : *La stratégie européenne pour les données*.

Les quatre grands piliers de la stratégie européenne pour les données

Les deux premiers piliers : les infrastructures et le partage de données

D'ici cinq ans, la quantité de données stockées et traitées au sein de l'UE sera multipliée par cinq. Elle sera ainsi passée de 33 zettabits en 2019 à près de 180 zettabits. Pour faire fonctionner l'IA, stocker et traiter les données dont elle a besoin, des infrastructures (le *brick and mortar*) sont indispensables pour interconnecter les centres de données – au sein d'un même État et entre États membres de l'Union européenne.

Le pendant à ces infrastructures est la création d'espaces de données européens communs, en particulier des espaces sectoriels. De fait, outre les échanges entre un industriel et ses fournisseurs ou ses clients, les acteurs concurrents d'un même secteur industriel peuvent avoir tout intérêt à partager leurs données pour des questions d'efficacité de marché. Mais il était important de mettre en œuvre les règles pour promouvoir la disponibilité des données et créer un environnement fiable pour faciliter leur utilisation ou réutilisation et la création de nouveaux services et produits innovants.

Troisième pilier : le marché du cloud et du edge computing

Alors que les infrastructures de données sont très centralisées, il convient d'encourager leur décentralisation afin d'amener les données et les capacités de traitement au plus proche des utilisateurs – l'objectif étant de permettre à ces derniers de renforcer leur pouvoir de contrôle et de décision sur l'utilisation et l'exploitation commerciale de ces données.

Pour encourager cette distribution des capacités de traitement, la Commission a mis pour objectif le déploiement de plus de 10 000 nœuds de edge computing (dans les hôpitaux, les centres de production et de logistique, les bâtiments, etc.) d'ici à 2030 sur l'ensemble de son territoire, afin que n'importe quelle entreprise ou administration publique ne se trouve jamais à plus de 100 kilomètres d'un nœud. Elle entend aussi encourager les utilisateurs à faire migrer leurs applications et données vers le cloud, lequel évite notamment d'avoir à investir dans des serveurs.

Depuis trois à cinq ans, le marché du cloud est marqué par un fort effet de concentration. Quatre gros acteurs mondiaux détiennent aujourd'hui les deux tiers du marché européen (et mondial) de l'infrastructure et des middlewares : Amazon Web Services, Microsoft Azure, Google Cloud et Alibaba Cloud. En parallèle, le nombre d'acteurs publics et privés migrant vers le cloud ne cesse d'augmenter. Aussi importe-t-il que les acteurs européens se saisissent de cette très belle opportunité pour se positionner sur le marché du cloud et sur celui, émergent, du edge computing.

Quatrième pilier : les textes réglementaires

Deux textes fondamentaux permettent et garantissent la circulation des données au sein de l'Union européenne : le règlement général sur la protection des données (RGPD) et celui instaurant la libre circulation des données non personnelles au sein de l'UE.

C'est en s'appuyant sur eux que la Commission s'efforce de construire un cadre réglementaire pour promouvoir l'ouverture des données – avec le Data Governance Act (DGA), l'Open Data Directive et le Data Act – et corriger les asymétries sur le marché de la donnée, avec le Digital Markets Act (DMA) et certains volets du Data Act.

Focus sur le Data Act

Le Data Act devra permettre un meilleur accès aux données générées par les objets connectés, afin de rendre aux utilisateurs de ces objets le pouvoir de décider comment et à quoi elles serviront.

La Commission entend résoudre les asymétries contractuelles qui peuvent exister entre un fournisseur de service de données ou de produits connectés et les utilisateurs. Il s'agit également d'octroyer aux autorités publiques la possibilité de rendre obligatoire le partage de données qui sont nécessaires à des fins spécifiques d'intérêt public.

La Commission a également décidé de créer un droit concernant le changement de fournisseur de services cloud, pour permettre aux utilisateurs de demander à n'importe quel moment le portage de leurs données et/ou applications vers un autre fournisseur équivalent.

En outre, un chapitre du Data Act vise à se prémunir contre le risque d'accès illégal à des données sur le territoire de l'Union européenne. En effet, nombre d'acteurs privés qui fournissent des services de données sur ce territoire sont originaires de pays tiers et sont donc soumis à d'autres législations que celles de l'Union européenne. Or ces législations peuvent leur imposer de donner accès à un certain nombre de données, y compris pour des infrastructures ou des services se trouvant sur le territoire européen. Et, très souvent, ces accès vont à l'encontre du RGPD et des autres législations européennes. En l'occurrence, le Data Act contraindra les fournisseurs de données opérant dans l'UE à mettre en œuvre toutes les mesures techniques, organisationnelles et légales nécessaires pour se prémunir contre ce genre de requête. Ce faisant, il renforcera la primauté du droit européen sur celui de pays tiers.

Quelques exemples de mise en pratique du Data Act

Le secteur de l'agriculture se développe beaucoup autour de l'économie de la donnée et les objets connectés déployés dans les exploitations agricoles sont de plus en plus nombreux, mais les agriculteurs n'ont pas forcément la possibilité d'accéder à toutes les données ainsi générées ni de les exploiter. Le Data Act permettra aux agriculteurs d'accéder plus facilement à leurs données, de décider avec qui ils souhaitent les partager et de révoquer l'accès à leurs données à tout moment.

Le secteur automobile représente un autre exemple. Les constructeurs automobiles ont la volonté de capter les données des utilisateurs pour les exploiter, en vue de vendre des services à valeur ajoutée. Là encore, le Data Act permettra à l'utilisateur d'avoir un libre accès aux données générées par son véhicule et de décider avec qui il souhaitera les partager.

En conclusion, la liberté de choix pour les utilisateurs constitue un des piliers fondamentaux du Data Act.

Commentaires

Jean-Yves Ollier, Conseiller d'État, Coordonnateur du n°18 de la revue Enjeux numériques

Le DGA, qui a été publié avant l'été et qui sera appliqué en septembre 2023, introduit une nouvelle figure dans l'économie des données : l'intermédiaire des données, qu'il distingue d'autres types d'acteurs. En outre, il fait évoluer les règles dans le sens d'un régime de neutralité, ce qui fait en partie écho à des logiques de séparation fonctionnelle déjà en vigueur dans d'autres industries. Pouvez-vous nous apporter des précisions sur ce qui a conduit à proposer ce modèle ?

Pierre Chastanet

Initialement réticents à cette pratique, les acteurs industriels ont fini par comprendre l'intérêt à mutualiser les données. La question de savoir qui allait gérer ces espaces communs de données s'est alors posée, et a fait émerger la notion d'intermédiaire et le concept de data broker. De fait, un *middle man*, acteur de confiance, est indispensable pour orchestrer ce partage de données. Il faut aussi que le système de mutualisation soit viable d'un point de vue économique, sachant que les données devront être anonymisées et agrégées en vue de leur exploitation. Un tel travail pourrait donner lieu à une rémunération.

Par ailleurs, si le Data Governance Act et le Data Act ont vocation à définir des règles généralistes communes à l'ensemble des secteurs économiques, rien n'empêchera pas de spécifier par la suite des règles de gouvernance et de partage permettant de répondre aux besoins de tel ou tel secteur.

Table ronde 1 : Les enjeux du partage des données

Modération : **Éric Brousseau** | Directeur scientifique, Chaire Gouvernance et Régulation & Club des Régulateurs, Université Paris Dauphine-PSL

Cette première table ronde est consacrée aux enjeux du partage des données à partir d'illustrations dans les secteurs de la santé, des transports, de l'agriculture, et de l'industrie de la production de données avec par ailleurs des approfondissements sur les modèles de rémunération des « producteurs » de données et sur l'enjeu des données pour entraîner les algorithmes d'IA.

Trois problématiques/logiques de partage des données doivent être distinguées :

- Tout d'abord, les données peuvent être partagées en vue d'assurer une certaine transparence dans la relation entre l'offre et la demande afin d'identifier des vecteurs d'amélioration, des outils, des processus, des systèmes et des arrangements organisationnels existants.
- Ensuite, les données partagées peuvent être le vecteur d'une intégration entre les prestataires et d'innovation en matière de service.
- Enfin, les données peuvent être partagées en vue d'entraîner des IA.

Ces différentes logiques ne portent pas les mêmes enjeux en matière d'adaptation de l'offre aux besoins, de potentialités dans la réalisation d'objectifs collectifs (notamment de politiques publiques), de compétition et partage de la valeur ajoutée entre opérateurs, de compétitivité internationale de ces derniers, ou encore en matière de sécurité individuelle et collective en cas d'utilisation malveillante des données. Les illustrations sectorielles qui suivent permettront de présenter les spécificités des enjeux propres aux domaines concernés, de présenter un certain nombre d'initiatives en cours, d'appréhender les difficultés de mise en œuvre rencontrées par les parties prenantes, et d'évoquer les besoins non couverts afin d'éclairer sur les perspectives souhaitables dans chacun de ces secteurs.

Les avancées en matière de partage des données de mobilité dans les transports

Jordan Cartier | Secrétaire général, Autorité de régulation des transports

Le secteur des transports est marqué par deux enjeux phares : le numérique et la transition écologique, dont les données de mobilité sont la clé de voûte. Ce n'est d'ailleurs pas sans raison que la Commission européenne a intitulé sa feuille de route pour les transports « Stratégie pour une mobilité durable et intelligente ».

L'ouverture des données de mobilité : un enjeu fondamental favorisant le développement d'applications numériques de mobilité et contribuant à la décarbonation des transports

L'ouverture des données de mobilité est la condition *sine qua non* pour le développement d'applications numériques de mobilité viables et efficaces, au bénéfice des usagers et de la collectivité. Pour les premiers, il s'agit de faciliter la planification des trajets, en particulier lorsqu'ils impliquent plusieurs modes de transport, mais aussi d'informer sur les horaires en temps réel, a fortiori en cas de perturbations. Pour la seconde, il s'agit de favoriser l'abandon de l'autosolisme au profit des transports collectifs, ce qui contribuerait à décongestionner les infrastructures routières, et d'atteindre ainsi l'objectif de décarbonation des transports fixé pour 2050, avec un premier jalon décisif en 2030.

Pourtant, l'accès aux données de mobilité ne va pas de soi. De fait, non seulement les producteurs de ces données sont nombreux et variés (autorités organisatrices des transports, entreprises locales, grands acteurs nationaux des transports), mais ils peuvent parfois craindre qu'une utilisation abusive ou discriminatoire soit faite de leurs données. Qui plus est, ces données peuvent s'avérer complexes à utiliser, de qualité variable voire incomplètes. Dans ce contexte, il est nécessaire de définir un cadre juridique.

Le cadre juridique européen et national vise à faciliter l'ouverture des données

L'ambition d'un tel cadre est de faciliter l'accès aux données via un point d'accès national unique, de les rendre interopérables et de garantir leur qualité.

Au niveau européen, le cadre juridique est fixé par la directive 2010/40/UE du 7 juillet 2010 sur les systèmes de transport intelligents et son règlement délégué (UE) 2017/1926 du 31 mai 2017. Il impose aux producteurs de données de les mettre à disposition dans un point d'accès unique et de respecter des standards de qualité et des formats déterminés pour rendre les données interopérables. Quant aux usagers de ces services, ils doivent utiliser les données de façon transparente et non discriminatoire, sans biais en faveur ou en défaveur d'un opérateur de transport en particulier.

Sur le plan national, la loi d'orientation des mobilités du 24 décembre 2019 a précisé et complété la réglementation européenne, et désigne le site www.transport.data.gouv.fr comme point d'accès unique aux données de mobilité.

Cette loi confie également au régulateur économique sectoriel des transports la régulation à la fois de l'ouverture des données numériques de mobilité et celle de la distribution numérique des titres de transport. Le législateur a ainsi chargé l'Autorité de régulation des transports (ART) d'évaluer et de contrôler l'ouverture et l'utilisation des données de mobilité, prévues par le règlement délégué et la LOM précités.

L'Autorité de Régulation des Transports, « gendarme » de l'ouverture des données de mobilité

L'action de l'ART en matière de contrôle de l'ouverture des données de mobilité s'articule autour d'un triptyque : « consulter, éclairer, contrôler. »

En ce qui concerne le volet « consulter », l’Autorité s’assure que les acteurs du secteur s’approprient bien les règles en vigueur et recueille leurs avis pour formuler, le cas échéant, des propositions d’évolution du cadre juridique. Tel était l’objet de la consultation publique organisée entre novembre 2021 et février 2022, qui a fait apparaître que le cadre réglementaire était connu, mais que son appropriation par les acteurs était difficile et que ses exigences n’étaient pas totalement maîtrisées.

En ce qui concerne ensuite le volet « éclairer », l’Autorité a publié, en juin 2022, un premier état des lieux de l’ouverture et de l’utilisation des données de mobilité. Par la suite, elle publiera un rapport sur une base au moins biennale.

Enfin, pour le volet « contrôler », l’Autorité a développé une méthodologie de contrôle, s’appuyant sur des outils de validation, qui lui permettra de vérifier le respect des règles en place par les producteurs et les utilisateurs de données. Elle dispose des pouvoirs nécessaires en matière de recherche et constat de manquement, ainsi que de poursuite en cas de manquement constaté. Son collègue peut décider de notifier des griefs et de saisir sa commission des sanctions le cas échéant. Le lancement d’une première campagne de contrôle vient d’être rendu public.

Les difficultés rencontrées appellent des évolutions de la réglementation

Ce cadre de régulation mérite à coup sûr d’évoluer, pour répondre à deux besoins : un besoin de clarification voire de simplification ; un besoin de coordination entre actions de régulation générale et actions de régulation sectorielle, notamment dans le cadre des nouvelles exigences particulières qui pourront être imposées par le DSA ou le DMA.

En conclusion

Les données de mobilité représentent l’un des deux piliers du secteur des transports. La billetterie multimodale est le second pilier des MaaS (*Mobility as a Service*), ces applications « couteaux suisse » qui doivent favoriser le report des utilisateurs vers des mobilités moins carbonées, en intégrant la dimension de l’information et celle de l’accès aux transports via la billetterie.

La loi d’orientation des mobilités a déjà préparé le terrain réglementaire, et les futurs textes européens pourront s’appuyer sur ce cadre national.

Les initiatives françaises et européennes en matière de gouvernance des données de santé

Clément Tonon | Auditeur au Conseil d'État

Un cadre de gouvernance robuste et efficace doit concilier plusieurs objectifs, d'autant que le sujet des données de santé a trait à la fois aux politiques régaliennes, à la question des valeurs, à la gouvernance, à l'aspect économique et à la sécurité.

Les principaux enjeux liés à la construction d'un cadre de gouvernance

Le cadre de gouvernance s'avère indispensable pour :

- améliorer l'usage primaire des données de santé pour les politiques de santé publique (soin et recherche) ;
- répondre à l'enjeu économique de l'usage secondaire des données de santé, dont la valeur devrait doubler d'ici à 2028, pour atteindre 50 milliards d'euros ;
- protéger les données personnelles, étant entendu que la donnée de santé brouille la frontière entre vie privée et enjeu économique et de marché ; cette ambiguïté se retrouve dans la double base légale du projet de règlement sur l'espace européen des données de santé, fondé à la fois sur l'article 14 du TFUE, qui a trait à l'organisation du marché intérieur, et sur l'article 16 du TFUE, qui concerne la protection des données personnelles ;
- répondre à la question de la valeur et du modèle de gouvernance, qui n'est pas univoque.

Quel modèle ?

Des différences notables s'observent selon les pays. Aux États-Unis, la loi HIPAA autorise la vente et le don de données de santé tant que le dossier de santé est dépersonnalisé. Le modèle scandinave, quant à lui, repose sur un système d'open data très abouti, avec des projets de partage de grandes bases de données et de valorisation via des techniques d'IA. Pour sa part, le Royaume-Uni a opté pour un modèle hybride, avec un acteur (HDR) chargé d'encadrer l'accès aux différentes bases et un secteur privé fortement associé par le biais de nombreux partenariats. Une dizaine d'acteurs privés ont ainsi un accès large et régulier aux bases de données.

Quant au modèle français, il se caractérise par un mouvement assez profond d'élargissement des bases, de leur concentration et d'ouverture de leur accès. Ce faisant, il suit le modèle de structuration du secteur sanitaire et social, très centralisé autour de l'Assurance Maladie, qui a très tôt favorisé l'élaboration de bases de données très qualitatives et a poussé le législateur à consolider autour d'elles d'autres bases plus petites, grâce à la loi de modernisation du système de santé adoptée en 2016. Cette démarche a abouti à la constitution d'une plateforme des données de Santé (*Health Data Hub*), qui met en relation une quarantaine de bases publiques et privées. Cette plateforme est ouverte aux opérateurs publics et, de plus en plus, aux opérateurs privés et aux praticiens.

Les limites du modèle national actuel sont assez bien résumées dans un rapport de 2022 consacré à l'analyse comparative des modèles de gouvernance de données de santé au sein de l'OCDE : la France est classée 4^{ème} sur 23 pays pour ce qui est de la robustesse de son modèle de gouvernance, mais seulement 14^{ème} pour la capacité de ce dernier à valoriser les données de santé au regard des différents enjeux précédemment évoqués. Cela s'explique par l'absence de bases dans certaines aires thérapeutiques (le cancer, le diabète, certaines maladies cardio-vasculaires) et le manque d'automatisation des bases existantes ; les difficultés temporelles pour collecter les données, les difficultés d'identification unique

entre les différentes bases et une faible utilisation des bases pour évaluer la performance du système global de santé ; les difficultés à consolider les différents systèmes de gestion de crise ; les limites dans la disponibilité et la qualité des données, notamment hospitalières.

Le projet d'espace européen des données de santé

Le 3 mai dernier, la Commission européenne a publié un projet de règlement relatif à l'espace européen des données, qui vise à encadrer l'utilisation de ces données au sein de l'UE, en particulier l'utilisation dite secondaire, grâce à la création de tiers de confiance, des organismes d'accès aux données de santé. Ceux-ci seront chargés de délivrer les autorisations d'accès au niveau national, garantir la transparence de cet accès, veiller au respect des objectifs et des modalités d'accès (publication des résultats, des travaux et de la démonstration de l'usage), ainsi qu'au respect des normes et des standards de cybersécurité.

L'avis du Comité Européen de Protection des Données (CEPD) sur ce projet de règlement souligne notamment la dichotomie entre l'exigence de protection des données de santé et leur valorisation économique. Relevant une inadéquation avec certaines dispositions du RGPD, il propose d'exclure du projet de règlement les données issues des objets connectés, essentiellement pour des raisons de protection de la vie privée. Il engage aussi le législateur à accentuer les contrôles d'accès exercés par les autorités nationales d'accès, en ajoutant un certain nombre de critères dont la démonstration d'un lien suffisant avec des enjeux de protection sociale et de santé publique. Enfin, le CEPD veillera à ce que le projet d'espace européen des données de santé n'aboutisse pas à la constitution d'une base de données de santé centralisée au niveau européen, et insiste pour que le stockage et le traitement des données de santé s'effectuent sur le territoire de l'Union européenne.

En conclusion

La donnée de santé, son encadrement et son exploitation soulèvent deux questions. D'une part, l'Union européenne peut-elle assumer cette ambition en termes de gouvernance sans une ambition industrielle concomitante ? D'autre part, peut-on faire confiance aux Etats européens pour s'accorder sur une vision et une projection commune en matière de données de santé ? Cette dernière question a été particulièrement exacerbée par la crise du Covid-19.

Les projets dans le secteur agricole et les enjeux industriels dans le contrôle des chaînes de production agricole et agroalimentaire, ainsi que sur les marchés des matières premières agricoles

Renaud Font | Directeur des opérations, Agdatahub

Agdatahub, premier intermédiaire de données

Porté par la profession agricole et l'implication assez forte des pouvoirs publics dans sa gouvernance, notamment via la Banque des Territoires de la Caisse des Dépôts, Agdatahub entend se positionner comme le premier intermédiaire de données, en tant que plateforme d'échanges et de consentement.

Les enjeux liés à cette ambition sont multiples car l'agriculture focalise de nombreuses attentes : une meilleure traçabilité et une plus grande transparence pour les consommateurs ; un meilleur suivi de production, une plus grande compréhension, une meilleure gestion et une amélioration de la qualité de vie au travail pour les agriculteurs.

Les données générées par les outils utilisés par les agriculteurs sont massives, très éclatées, très disparates et mises en silo par la pluralité des acteurs économiques (90 000 partenaires travaillent au quotidien auprès des 400 000 agriculteurs français et, à l'échelle européenne, ces partenaires se comptent en centaines de milliers). Or le RGPD s'avère inadapté au monde agricole, puisque 90 % des données produites par un agriculteur ne relèvent pas du registre de la donnée personnelle. Agdatahub entend saisir l'opportunité du DGA pour se positionner comme premier intermédiaire de données.

De l'importance de contrôler les données produites par les machines agricoles

La question du consentement

Les éleveurs de bovins et producteurs du lait sont désormais tous équipés en robots de traite, pour automatiser la collecte du lait. Quand les vaches passent à la machine à traire, des milliers de données sont collectées (même si l'agriculteur visualise seulement celles relatives au passage de la vache, au niveau de collecte du lait, à la quantité et à la détection d'une éventuelle maladie précoce). Mais, en cas de changement de matériel, l'historique de vie du cheptel est perdu. Pourtant, en cas de vente, la valeur d'une exploitation s'accroît de 25 % quand il existe des données historiques de production. Les moissonneuses-batteuses constituent un autre exemple frappant. Ces machines sont « truffées » de capteurs. Quand un agriculteur passe dans ses champs, le fabricant du matériel est informé avant lui de la qualité et du niveau de rendement des récoltes. Un fournisseur américain qui reçoit ces données peut les revendre sur les marchés, et ainsi influencer les cours des matières premières.

Dans tous les cas, ces données sont collectées sans le consentement de l'agriculteur. Ce problème devrait être résolu par le DGA et le futur Data Act, qui instituent le producteur agricole comme l'ayant droit des données produites par les machines qu'il utilise. Les fabricants des robots ne seront donc plus propriétaires de fait des données agricoles, et devront adresser une demande d'accès à ces données aux agriculteurs qui les produisent, préalablement à toute exploitation.

Les enjeux et les solutions envisagées

La question de la maîtrise de la collecte, du partage et de l'exploitation des données agricoles soulève plusieurs enjeux : un enjeu économique (la valeur d'une exploitation, l'influence des cours des marchés) ; un enjeu de compétences autour de la data ; un enjeu d'évolution culturelle pour faire prendre conscience de l'importance de la donnée.

En l'occurrence, Agdatahub offre des solutions d'infrastructures numériques disponibles pour l'ensemble des acteurs du secteur, permettant d'enregistrer et de faire circuler les données dans le respect du droit, l'enjeu final étant d'alimenter des algorithmes, l'IA et des modèles prédictifs.

De son côté, la Confédération nationale de l'élevage a mis en œuvre un véritable système sectoriel, afin d'inciter l'ensemble des acteurs économiques, parfois concurrents, à échanger des données au sein d'une infrastructure de confiance. Ce système fonctionne grâce à deux piliers structurants : l'échange de données et la validation de la conformité du contrat de licence d'accès aux données via l'interrogation du consentement de l'ayant droit.

En conclusion

Aujourd'hui, aucun outil ne permet aux agriculteurs de changer ou de mettre un terme aux contrats qui les lient aux fabricants qui récupèrent les données produites. L'Europe pourra-t-elle peser face aux géants du numérique, qui détiennent 75 % du marché et sur lesquels tout le monde s'est reposé jusqu'à présent ?

L'approche trans-sectorielle de l'analyse des enjeux à partir du rôle des données dans l'entraînement des algorithmes

Julien Chiaroni | Directeur du Grand Défi sur « l'IA de confiance pour l'industrie », Secrétariat Général à l'Investissement (SGPI)

L'importance de la donnée dans un contexte de complexité croissante des systèmes

Dans le secteur industriel, les applications seront de plus en plus complexes et portées sur le *edge*, pour développer des systèmes de plus en plus intelligents. Ces systèmes requerront de fortes compétences techniques, dans la mesure où ils seront sollicités pour du hardware, de la donnée, mais aussi de la connaissance métier, du cyber, de l'intelligence artificielle, etc. Cette évolution s'inscrit dans un contexte où la part du logiciel dans la valeur ajoutée des systèmes ne cesse de croître. À titre d'exemple, d'ici à 2030, 60 % de la valeur ajoutée d'un véhicule dépendra du logiciel qu'il contiendra (source : PwC).

La donnée est au cœur de l'IA, dont le récent rapport du think tank New Deal précise les différents champs techniques (source :). L'ensemble des technologies utilisées nécessiteront des données de qualité et toujours plus nombreuses pour atteindre les niveaux de performance exigés. La connaissance métier représentera aussi un élément fondamental, notamment dans l'optique de mécaniques d'hybridation.

La donnée servira à entraîner les systèmes, mais aussi à mener des tests et des évaluations. La partie *assessment* sera au cœur de la future réglementation européenne sur l'IA, ce qui imposera de respecter un certain nombre de valeurs et d'apporter de la confiance aux utilisateurs, c'est-à-dire de l'éthique, de la sûreté, de la transparence, etc.

Les aspects de qualité, de quantité et de couverture des données par rapport au domaine d'emploi constituent autant d'éléments primordiaux pour atteindre les niveaux de performance requis et garantir que le système ne fait que ce qu'on lui demande de faire. Les éléments de partage des données et de leur accessibilité pour le plus grand nombre sont aussi des plus importants pour atteindre ces objectifs.

L'implémentation

Comment faire en sorte que les exigences réglementaires et sociétales se traduisent par une mise en œuvre opérationnelle de la part de l'ensemble du tissu économique, à travers des produits applicatifs et des services ? Plusieurs actions peuvent être engagées pour relever ce défi : construire une gouvernance suffisamment cohérente et ouverte pour prendre en compte les attentes et les exigences de chacun ; définir des standards permettant de préciser les exigences, les spécifications, les caractéristiques, etc. (ce qu'a choisi de faire l'UE à travers sa nouvelle réglementation) ; prévoir les infrastructures techniques et les outils logiciels garantissant la gestion de l'ensemble des fonctionnalités et la déclinaison opérationnelle des exigences réglementaires à moindre coût ; trouver le juste équilibre entre innovation et régulation (le surcoût lié à la compliance est parfois estimé à 400 000 euros pour certains types d'application, en fonction de leur niveau de complexité).

L'interdisciplinarité et le tiers de confiance

L'aspect interdisciplinaire constitue un autre enjeu fondamental. Data scientists, spécialistes en IA, ingénieurs en logiciel, managers de la qualité et du risque, experts dans le domaine des systèmes, juristes, experts en sciences humaines et sociales... : la France a décidé d'investir massivement dans les compétences et la formation dans le domaine de l'IA, pour garantir une articulation cohérente entre la réglementation, la standardisation et les infrastructures techniques. Il convient également d'instaurer un dialogue étroit entre toutes ces compétences. De fait, il serait inutile de vouloir imposer certaines contraintes ou exigences que l'on serait dans l'incapacité d'implémenter techniquement.

À partir du moment où des contraintes et exigences réglementaires portant sur les données, le partage de données et l'IA seront imposées, il sera fondamental de disposer de compétences et d'organismes spécialisés dans l'audit et l'*assessment* pour instaurer de la confiance. L'homologation d'un véhicule autonome, par exemple, n'aura rien de comparable avec celle d'un véhicule actuel. Malheureusement, l'auditeur, le certificateur ou l'homologateur qui représente un tiers de confiance reste encore trop souvent le « parent oublié » des réflexions.

Avantages et inconvénients des modèles alternatifs de rémunération du partage

Denis Berthault | Président, Groupement français de l'industrie de
l'information (GFII)

L'indispensable qualité des données

Les différents types d'outils d'intelligence artificielle ou le metavers ont besoin d'être alimentés par d'énormes volumes de données de grande qualité pour tester ou valider des processus ou solutionner un maximum de cas de figure (« use case »). Dans ce contexte technologique de pointe, comment cerner le concept de qualité appliqué à des données ? Il peut être décliné ainsi : exhaustivité, fraîcheur, temps réel, mise à disposition par API robustes et pérennes, sécurité de l'hébergement, niveau d'anonymisation adapté à l'usage, stockage « froid » ou « chaud ».... Pour être opérante, il est également nécessaire que cette qualité soit constante, vérifiée et contrôlée en permanence afin d'éviter l'envoi de données biaisées (incomplètes, non exhaustives, obsolètes). Ce n'est qu'une fois cette qualité constatée, que le réutilisateur des données pourra développer, en toute confiance, de nouveaux services ou applications.

La question sous-jacente du coût

Ce niveau de qualité suppose d'importants moyens financiers, une équipe, une gouvernance, des process et un niveau d'expertise qui représentent un investissement initial significatif et récurrent. Tous les acteurs professionnels de la chaîne de valeur de la donnée en ont fait l'expérience. Il naît une contradiction si, face à ces investissements, le fournisseur de données fournit ses données gratuitement, dans le cadre d'une licence dite ouverte (ce qui est l'acception du terme « open data » en France. Comment, dans ces conditions, identifier le retour sur investissement quand le producteur s'interdit de connaître les organisations réutilisant ses données, à quelles fins elles sont utilisées ou déterminer le chiffre d'affaires incrémental réalisé grâce à elles ? Il ne fait pas compter sur les réutilisateurs pour déclarer spontanément ce qui relève quasiment du secret industriel : une entreprise qui agrège des données A et B gratuites pour créer et vendre un produit C payant, sera peu encline à divulguer le *modus operandi* par lequel elle amalgame et enrichit ces données, pas plus que les fournisseurs auprès desquels elle s'est approvisionnée.

En conclusion

Il n'existe donc pas de donnée exploitable dans un contexte professionnel sans qualité, il n'existe pas de qualité sans investissement et il n'existe pas d'investissement sur le long terme sans un retour sur investissement mesurable. Ceci posé, il n'est pas certain que les acteurs publics - contraints à une mise à disposition gratuite – parviennent à long terme à suivre la demande exponentielle de données destinées à nourrir IA ou metavers, tant en quantité et en qualité. Cela pourrait-il déboucher à court ou moyen terme sur la désintermédiation de certains acteurs publics par des acteurs privés, pas forcément français ? On n'en jurerait pas.....

Dans ce contexte pessimiste, il est urgent de s'inspirer du prochain Data Governance Act européen⁽²⁾, lequel ouvre de nombreuses possibilités aux acteurs publics en leur permettant, entre autres, d'établir une redevance pour l'utilisation de données publiques dites protégées, *a fortiori* quand leur mise à disposition aura nécessité un effort particulier (anonymisation, protection du secret des affaires ou des secrets d'entreprise, droit de propriété intellectuelle...).

- (1) La mission présidée par Éric Bothorel a remis son rapport au Premier ministre le 23 décembre 2020 <https://www.vie-publique.fr/sites/default/files/rapport/pdf/277879.pdf>
- (2) Le règlement UE 2022/868 sur la gouvernance des données, publié au Journal officiel de l'Union européenne le 30 mai 2022, entrera (comme le RGPD avant lui) automatiquement en vigueur en septembre 2023. Selon la Commission européenne, il pourrait générer de 7 à 11 milliards d'euros par an rien qu'au titre de l'échange de données <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:32022R0868&from=FR>

Conclusion des intervenants

Jordan Cartier

Les problématiques et les enjeux afférents au partage des données sont globalement partagés. Le secteur des transports paraît assez en avance en la matière, peut-être parce que la directive 2010/40/UE du 7 juillet 2010 sur les systèmes de transport intelligents avait anticipé l'importance des services d'information multimodale. En outre, le point d'accès national (PAN) a permis de pallier la difficulté d'éclatement des données et la puissance publique a mis en place des standards et un régulateur pour s'assurer de leur respect par les producteurs.

Concernant le coût de production de données de qualité, la loi d'orientation des mobilités et les textes réglementaires d'application contiennent des dispositions prévoyant que des compensations financières peuvent, lorsque l'utilisateur sollicite le service de fourniture de données au-delà de certains seuils, être versées aux producteurs de données dynamiques, sous le contrôle du régulateur.

Clément Tonon

Les différentes interventions montrent l'intérêt d'adopter une approche comparative. Celle-ci sera d'ailleurs de plus en plus nécessaire au fur et à mesure que la Commission émettra ses textes réglementaires sectoriels, pour mieux appréhender les rapprochements possibles et les différences. Nous nous trouvons encore « au milieu du gué », dans la mesure où seuls les textes transversaux sont publiés. Il sera sans doute nécessaire de nous revoir dans quelques mois pour dresser un bilan et envisager les possibilités en termes d'implémentation et de concrétisation des projets.

Renaud Font

Je retiens les convergences et l'urgence de définir un cadre réglementaire au niveau européen. Pour le secteur agricole, il est fondamental de sécuriser contractuellement l'identification et les principes d'accord à l'accès aux données et l'échange de données, mais aussi de développer des solutions techniques et des infrastructures permettant de déployer des solutions de type edge au plus près du développement des usages.

Julien Chiaroni

Pour suivre une approche offensive en termes de réglementation et de régulation, il est essentiel de ne pas négliger les aspects de politique industrielle et de stratégie d'innovation, car nous faisons face à des marchés et des modèles très complexes, qui requièrent un soutien et un financement parfaitement adaptés.

Denis Berthault

Depuis quelques années, le domaine de la donnée est de plus en plus encadré juridiquement, tant au niveau européen que français. L'implication du gf2i dans la préparation du Règlement européen sur la gouvernance des données lui a permis de constater que le processus législatif et réglementaire européen, très élaboré et ouvert à tous, apportait une technicité et une précision qui faisait parfois défaut aux textes français, ce qui, dans un domaine aussi technique que les données, débouchait sur des résultats mitigés.

Table ronde 2 : Comment organiser le partage des données ?

Animation : **Jean-Yves Ollier** | Conseiller d'État, Coordonnateur du n°18 de la revue *Enjeux numériques*

La première table ronde nous a montré que les caractéristiques propres à chaque secteur économique (acteurs, chaîne de valeur, défaillances de marché, objectifs d'intérêt commun) et aux données qui y sont associées (personnelles/non personnelles, sensibilité, cycle de vie, contrôle d'accès) dessinaient des architectures et des régimes de partage différents, mais aussi des problématiques communes. L'objet de cette seconde table ronde est de débattre des questions de mise en œuvre communes aux différents écosystèmes de partage et aux espaces européens de données.

Elle portera, d'une part, sur les outils et les infrastructures, d'autre part, sur l'organisation et la gouvernance du partage. Il sera aussi question de la manière dont se fait l'articulation entre ces deux composantes.

Les principales fonctions des opérateurs de plateformes et fournisseurs de services de partage de données. Quels sont les changements portés par le DGA ?

Laurent Lafaye | Co-CEO, Dawex

La stratégie européenne de la donnée

La notion de plateforme d'échange de données

En 2020, la Commission européenne a défini une stratégie européenne pour les données dans la continuité de ce que la France avait identifié dès 2018 après le rapport Villani sur l'intelligence artificielle, en l'occurrence le fait que l'Europe n'aura pas de leadership sur l'IA tant que le monde académique et les entreprises ne pourront accéder aux données, y compris celles couvertes par le secret des affaires.

La Commission entend donc voir émerger des espaces de données, en définissant un certain nombre de règles, à la fois techniques et contractuelles ; associer aux échanges de données un volet économique, dans la mesure du possible ; assurer une conformité aux différentes réglementations relatives à la circulation des données, que ces réglementations soient déjà en place, nouvelles ou à venir (notamment les réglementations sectorielles).

Même si elle n'est pas nouvelle, cette notion de plateforme d'échange de données est relativement naissante – en témoigne l'organisation des agriculteurs français autour de la plateforme Agdatahub.

La problématique de la circulation de la donnée

Collecte, stockage, traitement, circulation de la donnée : chacun fait l'objet de cadres technologiques spécifiques et de réglementations nouvelles. La collecte et le stockage de la donnée constituent un élément clé de la confiance dans l'accès et l'échange de données. À ce sujet, la maîtrise de la territorialité de la donnée pose question. Les États-Unis ont d'ailleurs défini un cadre d'accès aux données, y compris lorsqu'elles sont produites et collectées par des entreprises européennes. En Europe, le cadre réglementaire se structure progressivement.

L'instauration d'un cadre de confiance à l'échelon national, européen et transnational

Identifier les acteurs

Le sujet de la circulation est souvent mis de côté dans la chaîne de valeur de la donnée. La stratégie européenne pour les données entend y remédier, en s'appuyant sur le Data Governance Act (DGA) et le futur Data Act.

Le DGA définit le rôle des opérateurs d'écosystèmes ou d'espaces de données, qui était jusqu'à présent peu identifié ou peu compris, afin de créer un cadre de confiance pour les autres acteurs de l'économie de la donnée. Un parallèle est souvent dressé avec les marchés financiers avec d'un côté ceux qui construisent des produits financiers et de l'autre, les plateformes qui permettent de les échanger.

Par ailleurs, l'identification des personnes morales – le producteur et l'acquéreur de la donnée – représente un élément connexe mais incontournable dans l'instauration de la confiance pour l'échange de données. De même, le système qui permet d'opérer ces échanges doit être identifié comme un état de confiance. Dans cette optique, le DGA introduit la notion de certification des fournisseurs de service d'intermédiation de la donnée. Il conviendra de déterminer rapidement quelle autorité pourra délivrer cette certification.

Contractualiser les échanges de données

Le sujet de la contractualisation de l'échange de données est également nouveau. Il doit se traiter à travers le Data Act, futur règlement européen qui introduira des modèles de licence.

Le Japon a signé des accords d'adéquation sur la donnée avec l'Union européenne dès 2019 et suit de près les travaux actuellement menés au sein de Gaia-X. La question d'interopérabilité des espaces de confiance, qu'ils soient sectoriels ou territoriaux, représente aussi un élément clé dans le développement de l'économie de la donnée.

En conclusion

Le fait que d'autres zones du monde s'intéressent à ce que construit l'Europe en matière d'accès, d'échange et de distribution des données est important pour les entrepreneurs, car cela démontre que leurs technologies pourront s'exporter et que la confiance pourra s'instaurer bien au-delà des frontières de leur pays et même de l'Europe.

Le développement de la gouvernance de l'infrastructure de cloud et les éléments critiques pour la promotion du partage des données

Anne-Sophie Taillandier | Directrice de TerraLab, Institut Mines-Télécom
& **Alban Schmutz** | Co-fondateur de CISPE (Cloud Infrastructure Services Providers in Europe)

L'organisation du partage de données recouvre les infrastructures techniques, les échanges de données et la gouvernance. Pour partager des données, il est nécessaire de les stocker dans un data center qui comprend des serveurs et une infrastructure logicielle propre à les exploiter. En l'occurrence, la possibilité de récupérer ses données stockées chez un opérateur constitue l'un des premiers freins au partage de données. Comment avoir une vision « souveraine » de la gestion de ses données, y compris au niveau de l'infrastructure ? Pour répondre à cette question, il est intéressant de se pencher sur l'impact des licences de logiciel.

Pour un meilleur fonctionnement du marché

L'association CISPE a travaillé à développer, en collaboration avec le CIGREF, dix principes pour définir une licence logicielle équitable. Aujourd'hui, les éditeurs de logiciels imposent leurs conditions aux utilisateurs pour les empêcher de récupérer leurs données et leurs licences et aller « voir ailleurs ». L'échec de l'autorégulation dans le cadre du règlement sur la libre circulation des données non personnelles, évoqué par Pierre Chastanet, est très révélateur : plusieurs éditeurs de logiciels ont effectué un véritable « travail de sape » pour éviter que ce projet n'aboutisse.

Le Digital Markets Act (DMA), bien que fondateur, est insuffisant pour résoudre ce problème. Certains éditeurs adoptent des comportements de prédation vis-à-vis non seulement des parts de marché de leurs concurrents, mais aussi des données des utilisateurs, ce qui a donné lieu à un certain nombre de plaintes déposées auprès de la Commission européenne et des autorités de la concurrence au cours des 18 derniers mois.

Par ailleurs, le sujet porté par CISPE ne se limite pas aux frontières de l'Union Européenne. En effet, le travail réalisé et les principes ainsi définis ont été récemment repris à leur compte par différents acteurs en Israël, en Afrique du Sud et aux États-Unis.

La démarche Gaia-X

Les acteurs en position dominante sur le marché européen imposent leur dictat. De ce fait, le marché ne fonctionne pas correctement. La démarche Gaia X vise à régler ces dysfonctionnements en impliquant les acteurs (utilisateurs et fournisseurs) et en aiguillant le régulateur vers ce qu'il conviendrait de faire. L'objectif est d'écarter le risque de *vendor lock-in* (fait de rester prisonnier d'une solution d'un fournisseur).

L'association industrielle Gaia-X a été créée par 22 membres fondateurs, pour moitié français et allemands. Elle existe de façon officielle depuis janvier 2021 et compte 360 membres, représentatifs de l'ensemble du tissu économique européen. Les *cloud service providers*, les utilisateurs de ces clouds et les hyperscalers en font notamment partie.

Gaia-X cherche à proposer des solutions d'implémentation et de régulation, ainsi qu'un label à trois niveaux destiné à rendre les offres de cloud comparables : un niveau 1 portant sur la transparence ; un niveau 2 relatif à l'obligation de détenir des données stockées et traitées en Europe ; un niveau 3 portant sur l'immunité par rapport à des lois territoriales non-européennes. Elle entend également définir des règles d'interopérabilité pour favoriser la circulation et le partage de données, grâce au développement de

nouveaux produits et services sur différents espaces de données et dans différentes thématiques. Enfin, elle a vocation à pousser les documents fondamentaux qui intègrent les architectures techniques, les spécifications techniques et les labels précités (comprenant 62 critères), en vue d'instaurer la confiance nécessaire à une bonne circulation des données.

Gaia-X développe ses propres spécifications techniques, lesquelles pourront être compatibles à la fois avec les produits des éditeurs de logiciels et avec l'open source.

Les leçons à tirer du modèle de gouvernance du logiciel libre pour l'organisation du partage de données

Robert Viseur | Chargé de cours et chef de service TIC (Technologies de l'Information et de la Communication), Université de Mons, Belgique

Les stratégies de logiciels libres (open source) s'inscrivent dans le cadre plus large des stratégies d'ouverture. Trois types de stratégie peuvent être mobilisées pour valoriser un bien : les stratégies privées et propriétaires ; les stratégies d'open source ; les stratégies hybrides.

Le principe de l'open source

L'open source comprend des projets plus ouverts que d'autres. On peut y trouver un système de règles (liées à la propriété intellectuelle) ou des stratégies de déploiement de ressources pour essayer de contrôler un projet. Un parallèle peut être dressé avec les domaines des données (open data et données ouvertes d'intérêt public), du matériel (open hardware) ou du stockage (open cloud).

Parfois présenté comme une négation du droit d'auteur, le système d'open source ou de licence libre est en réalité une utilisation du droit d'auteur pour organiser le partage d'une ressource. Aujourd'hui, une dizaine de licences libres très populaires sont utilisées, avec trois ou quatre modalités typiques. Elles offrent un accès très large aux ressources. On retrouve l'équivalent dans le domaine des données, avec les licences utilisées pour protéger des bases de données, auxquelles sont incorporés certains cadres juridiques spécifiques, comme le droit *sui generis* des bases de données en Europe.

L'intérêt de recourir aux logiciels libres

En publiant des logiciels libres, les acteurs se privent d'une source de revenus – sur laquelle repose le modèle économique d'entreprises comme Adobe, Oracle et Microsoft.

Au-delà d'une possible position idéologique, il importe de comprendre que le fait d'abandonner la propriété du logiciel n'empêche pas sa valorisation. Dans le domaine du logiciel, en effet, une large part des revenus provient des services (assurance, assistance, adaptation). Ces différents éléments de proposition de valeur peuvent se retrouver avec la donnée : garantir la fiabilité et un niveau durable de qualité de la donnée, prévoir une aide à l'utilisation, etc. Le concept de logiciel ou de licence libre est donc tout à fait compatible avec des modèles d'affaires commerciaux.

L'intérêt d'un logiciel libre provient certes de son logiciel stock, c'est-à-dire le code source auquel tout le monde a accès – mais aussi, et surtout, de son logiciel flux –, autrement dit toute l'activité contributive qui peut être suscitée autour du logiciel.

En outre, la gouvernance de projet libre intègre la licence d'open source, qui définit un certain nombre de modalités d'accès et d'utilisation du logiciel. Elle introduit plus ou moins de règles en fonction de la notoriété, du succès, de l'importance du logiciel libre et de la diversité des acteurs qui y participent. Une architecture de participation permet de maximiser les contributions à un projet dans un cadre plus ou moins ouvert. Certaines logiques de gouvernance sont davantage communautaires et très ouvertes. D'autres, de type consortium, sont plus fermées. C'est notamment le cas d'Android.

L'open data

Dans un certain nombre de cas, le coût de production de l'open data est élevé, mais l'activité de coproduction de la donnée permet d'abaisser ce coût. Certains projets communautaires ont ainsi parfois été rejoints par des entreprises, étant entendu que les services qui se fondent sur les données libres peuvent donner lieu à des stratégies de valorisation plus classiques. Open Street Map en constitue un bel exemple.

Le contexte de l'open data est particulier, dans le sens où certains organismes, notamment publics, peuvent produire des données de référence. C'est le cas du SPF Santé qui publie des données de référence fiables et de qualité au sujet du Covid-19. L'enjeu de création d'un flux de contribution se trouve alors reporté au niveau de la réutilisation de ces données.

En conclusion

Il importe de bien comprendre que les stratégies d'open source sont intéressantes en ce qu'elles suscitent des activités collaboratives propres à créer de la valeur.

L'organisation du partage entre les différentes parties prenantes et les enjeux éthiques

Claude Kirchner | Directeur de recherche émérite, INRIA (Institut national de recherche en science et technologies du numérique)

Bien s'entendre sur la notion de donnée

Avec la conversion au numérique, le monde change sa façon de penser. Dans ce contexte, le terme de « donnée » est-il toujours approprié ? Les données sont-elles effectivement un « don » ? Rien n'est moins sûr ! Il est davantage question d'information que de don, dans un contexte de conversion numérique. Or le vocabulaire a une influence très forte sur notre façon d'appréhender l'objet information. Certains pourraient considérer qu'une donnée n'a pas de valeur en soi et qu'on peut donc la donner...

Une information peut être textuelle ou numérique, ou encore un programme, lequel correspond à une donnée complexe. Certains projets « collectionnent » les programmes. Le meilleur exemple est Software Heritage, qui vise à rendre accessible l'ensemble des logiciels créés par l'homme au cours des soixante-dix dernières années.

L'internet des objets

L'internet des objets, qui connaît une croissance très importante, induit une utilisation massive de données de différentes sortes : des données de communication complexes, qui font intervenir les opérateurs et les entités de dialogue ; des données autour de ce qui est produit et échangé ; tout ce qui contribue à l'exploitation de ces données, qui peuvent servir pour de l'apprentissage, des tests, de la validation ou encore des audits. Prenons quelques exemples issus de nos travaux avec France-Stratégie ou dans le cadre du Comité national pilote d'éthique du numérique.

Les véhicules dits autonomes

Les véhicules à conduite automatisée font partie des objets connectés mobiles dits « autonomes ». Ils dépendent d'architectures fiables, sécurisées et à faible latence. Ce type d'objets pose des enjeux d'éthique environnementaux, de relations de travail et de filière pouvant amener à réguler et légiférer.

Les agents conversationnels

Les agents conversationnels constituent un autre exemple intéressant d'objets connectés. Par exemple l'enceinte connectée Alexa a été vendue à 100 millions d'unités en 2018. Cet agent conversationnel inclut différents protocoles spécifiques pour la commande d'autres objets, comme les volets roulants, le chauffage, etc. – que l'utilisateur ne peut donc alors plus choisir librement. En l'absence d'une vision maîtrisée de ce type de déploiement, le risque existe d'aboutir à des situations de monopole inacceptables.

Autre point délicat, certains chatbots vont même jusqu'à simuler le discours et éventuellement l'image d'une personne décédée, en s'appuyant sur des données réelles pour en synthétiser de nouvelles.

La réalité étendue et le métavers

Ce type d'applications posera de nouveau la question des données personnelles acquises et passées à travers un crible. Un casque de réalité virtuelle, par exemple, peut inclure jusqu'à quatre ou six caméras. Les données acquises par ces caméras sont intégrées dynamiquement aux informations présentées via le casque à l'utilisateur, ce qui lui permet par exemple de « voir au travers » du casque. Le système de traitement de l'information capture tout en temps réel, y compris les réactions de l'utilisateur. Ce type d'objet connecté est donc particulièrement intrusif relativement aux données personnelles.

Les innovations dans ce domaine sont impressionnantes, mais posent la question de la gestion des données, fournies et acquises de façon très dynamique et quasiment sans aucun temps de latence.

Échanges

32

Daniel Berthault

Le modèle des licences de *share-alike*, recommandé par le secteur public, me semble très adapté à un usage amateur ou grand public, mais pas pour un usage professionnel. En effet, il n'apporte aucune garantie de qualité ou de fréquence des mises à jour.

Robert Viseur

La qualité d'un logiciel dépend d'un ensemble de pratiques liées à des méthodologies de développement et à des outils d'industrialisation des développements, qui permettent d'en améliorer la qualité. Le fait que des communautés de grande taille utilisent un logiciel open source permet aussi de détecter les bugs et de les corriger au fur et à mesure, ce qui aboutit à des logiciels de bonne qualité, tels que Linux, Firefox et Libre Office. La qualité n'est pas une propriété de base. Il existe des logiciels libres de très bonne qualité et d'autres de très mauvaise qualité. Il en va de même pour les données. En comparant les bases de données de Wikipédia à celles d'experts, j'ai constaté un bon niveau de qualité. Sur Wikipédia, en plus des vérifications humaines, de nombreux robots analysent aujourd'hui les contenus publiés et contribuent ainsi à en améliorer la qualité.

Pour qu'un logiciel, une fois installé, ou un jeu de données, une fois utilisé, conservent durablement leurs qualités, et leur valeur, un travail d'intégration des mises à jour est nécessaire. Un stock de données peut ainsi être mis en ligne en mode *share-alike*, puis l'activité de mise à jour de ce stock peut être facturée.

Daniel Berthault

Je ne nie pas l'intérêt que représente le *share-alike* dans le domaine des logiciels. Mais, une fois envoyée, une donnée ne revient pas. Elle peut être modifiée ou corrigée après coup, mais une fois envoyée, elle est « digérée », agrégée, et donne lieu à des décisions prises, notamment, par des robots. Il ne serait pas possible de confier à un fournisseur de données aléatoires la sécurité d'un véhicule autonome, par exemple.

Je reconnais également l'intérêt d'une application comme Wikipédia, mais celle-ci n'est liée à aucune prise de décision stratégique. C'est pourquoi je considère que le *share-alike* n'est pas adapté à un usage professionnel. C'est aussi la raison pour laquelle la révision des conditions de licence sera probablement envisagée dans la phase 2 de l'open data en France.

Robert Viseur

Deux projets co-existent avec Wikipédia : Wikidata, qui permet de gérer les données quantitatives associées à l'encyclopédie, et DBpédia, qui est la version rdf de l'encyclopédie et qui a vocation à être consommée par des machines. Dans la pratique, Wikipédia n'est donc pas qu'une application destinée aux amateurs ou au grand public. Les règles, les contrôles et les contributions sont telles que le niveau d'exigence de qualité de cette application ne cesse de s'accroître.

Robert Viseur

Deux projets co-existent avec Wikipédia : Wikidata, qui permet de gérer les données quantitatives associées à l'encyclopédie, et DBpédia, qui est la version RDF de l'encyclopédie et qui a vocation à être consommée par des machines. Dans la pratique, Wikipédia n'est donc pas qu'une application destinée aux amateurs ou au grand public. Les règles, les contrôles et les contributions sont telles que le niveau d'exigence de qualité de cette application ne cesse de s'accroître.

Anne-Sophie Taillandier

Le partage de la donnée n'est pas un acte humaniste, car il s'accompagne toujours d'une activité commerciale. Un propriétaire de données définit toujours une politique d'usage de ses données.

Gaia-X travaille sur ces sujets du contrôle de l'usage de la donnée, en essayant de définir les spécifications techniques et juridiques. L'intérêt est de trouver le moyen de rendre interopérables les différentes façons de fonctionner des écosystèmes. La notion de communauté est très importante. C'est pourquoi les hyperscalers, qui sont majeurs sur le marché du cloud et vont de plus en plus loin dans la chaîne de valeur, font partie des membres de Gaia-X. L'interopérabilité doit se définir avec eux aussi.

Denis Berthault

La notion d'open data montre ses limites dans l'absence de contrôle qu'elle implique.

Jean-Yves Ollier

L'open data, avec son degré d'ouverture maximale, se situe à l'une des extrémités du spectre des différentes modalités de mise à disposition des données, avec les licences qui y sont associées.

Denis Berthault

Quand le Conseil d'État met à disposition ses données de jurisprudence, il « habille » son contrat d'une licence type et il a bien raison car ces données sont délicates à manier. Le contrat standard initial a des limites et quand les données deviennent complexes ou sensibles, il faut l'adapter.

Robert Viseur

Quand la licence EUPL (licence de logiciel libre européenne) est sortie, une entreprise a par exemple estimé qu'il manquait certaines clauses. Elle a décidé de rajouter une clause de publicité. Une autre clause a été ajoutée pour préciser que l'utilisation de ce logiciel ne pouvait avoir lieu que dans un cadre gratuit. Ces ajouts ont eu deux conséquences : créer une insécurité juridique importante et un frein à la réutilisation du logiciel. Il semble donc préférable de créer une licence spécifique dès le début, plutôt qu'ajouter a posteriori des clauses à une licence standard, qui pourrait en altérer les caractéristiques connues.

Jean-Yves Ollier

Pierre Chastanet a fait référence à l'échec de l'autorégulation dans l'application des règlements sur les échanges de données. Aujourd'hui, une forme d'autorégulation fédérée se construit dans le cadre de Gaia-X. De son côté, Julien Chiaroni a évoqué l'importance, dans le projet de règlement sur l'IA, du renvoi à la normalisation, qui est un mode de co-régulation, associant différents acteurs économiques sous l'égide de l'autorité publique et dont l'équilibre en termes de gouvernance peut être fragile. Quels commentaires vous inspire l'articulation entre autorégulation et normalisation ?

Alban Schmutz

Plusieurs problèmes se sont posés avec l'autorégulation en matière de libre circulation des données non personnelles. Le règlement afférent à la libre circulation des données non personnelles prévoit, dans son article 6, que la portabilité de ces données puisse donner lieu à des codes de conduite sectoriels développés par l'industrie. Or les groupes de travail constitués pour développer ces codes de conduite ont été chapeautés par la Commission européenne, laquelle a émis des souhaits quant à la manière de gérer les choses. Celle-ci n'a donc pas laissé les acteurs économiques s'autoréguler.

J'ai co-piloté, avec le représentant des utilisateurs de l'association européenne des directeurs informatiques EuroCIO (dont le CIGREF est membre), un groupe de travail dédié à la rédaction d'un code de conduite sur la Portabilité des données au niveau des infrastructures cloud (IaaS – Infrastructure as a Service). À partir du moment où certains acteurs (Microsoft et SAP) sont entrés dans la salle, ils ont mis un coup d'arrêt à notre travail ! J'avais anticipé cet obstacle, en travaillant très rapidement et dès le début sur le fond du sujet. En parallèle, un autre code de conduite portant lui sur le niveau des applications logicielles (SaaS – Software as a Service) a été développé par d'autres acteurs, qui se sont focalisés sur les sujets relevant de la gouvernance, mais jamais sur le fond.

La Commission Européenne a imposé le fait que les deux codes soient publiés en même temps. Ce processus a duré trois ans avant que quelque chose ne se passe. Les évolutions récentes sont les plus intéressantes : une structure juridique a été mise en place pour faire appliquer ces codes de conduite. Il a notamment été décidé, au sein de cette structure, que les personnes qui participaient à la gouvernance de l'organisation devraient la quitter, si jamais elles ne déclaraient pas des services conformes aux outils d'autorégulation définis. Microsoft a alors saisi les autorités européennes de concurrence, en estimant que cette disposition représentait une atteinte à ses droits sur le marché, mais a finalement été débouté. Microsoft a quand même obtenu de cette façon un sursis de quatre ans.

L'échec de l'autorégulation est donc, en partie, l'effet de l'ingérence de l'autorité administrative de référence dans le processus. Par conséquent, il serait erroné de croire que l'autorégulation ne fonctionne pas. En revanche, il importe de déterminer la manière dont on met en place cette autorégulation. Cette question en soulève d'autres, notamment celles de la prédation de la valeur économique et d'entrisme dans la gestion de la gouvernance du système. Il importe de bien comprendre les intérêts de cette prédation pour pouvoir les contrebalancer par la gouvernance.

Quand nous avons constitué Gaia-X, nous nous sommes farouchement battus pour faire en sorte que sa gouvernance ne représente que les intérêts européens, d'où la décision de n'y inclure que des structures dont le siège social est basé en Europe au sein de son Conseil d'Administration.

Le sujet de la géopolitique est central. Je suis assez impressionné par les progrès réalisés en droit européen au cours des cinq dernières années, mais devons encore muscler notre jeu car certains éditeurs de logiciels ont œuvré pour sortir de la classification des « gatekeepers » du DMA pour échapper à certaines obligations et spécifications. En outre, certaines propositions de modification du texte dans le cadre du processus législatif européen et portés par des responsables politiques de premier plan au Parlement européen ont été écrites par des géants du secteur comme Microsoft ou SAP.

Anne-Sophie Taillandier

Ces enjeux sont essentiels pour l'autonomie et la stratégie que nous souhaitons atteindre au niveau européen.

Claude Kirchner

Le témoignage d'Alban Schmutz démontre bien la complexité des enjeux stratégiques et géopolitiques. Je le mets en lien avec la notion de qualité précédemment évoquée. Dans un certain nombre de comités, comme celui de l'AFNOR, on essaie de comprendre comment il est possible de préciser le sens de la qualité, ce qui nécessite d'impliquer tous les acteurs depuis ceux de la recherche jusqu'aux fournisseurs d'infrastructures. Ce processus d'élaboration de standards et de normes est très prenant. Il faut savoir que nos collègues européens, chinois ou américains prennent part à ces comités en ayant défini une stratégie très élaborée, ce qui n'est pas toujours le cas des Français. Il est donc important de reconnaître l'importance de s'engager dans ces processus.

Alban Schmutz

Nous avons effectivement constaté une forte différence entre l'organisation française et l'organisation allemande, au sein de Gaia-X.

Robert Viseur

Dans la perspective des stratégies d'ouverture, il importe de considérer la qualité comme un objectif qui peut être atteint par itérations successives. Linux en est l'illustration. Les open data Covid de SPF Santé étaient au départ très embryonnaires, mais ont pu être améliorées en l'espace de deux ans grâce aux *feedbacks* reçus et à un enrichissement progressif.

Eric Brousseau

Comme le disait Lawrence Lessig, il y a une vingtaine d'années, avec son fameux adage « Code is Law », les usages des biens informationnels dans le domaine du numérique dépendent autant des normes légales que des normes techniques qui s'imposent aux formats de données, aux algorithmes, etc. Il faut donc ne pas s'en tenir aux seuls principes légaux, mais analyser finement comment ils sont implémentés et rendus applicables ou non applicables sur le plan technique par les normes qui s'imposent ; d'où l'importance des questions de normalisation technique.

Intervention de clôture

Luc Rousseau | Vice-président du Conseil général de l'économie, de l'industrie, de l'énergie et des technologies au Ministère de l'économie, des finances et de la souveraineté industrielle et numérique

L'économie s'est longtemps organisée sur un axe « capital et travail » ; un second axe, « énergie et machines », a été à l'origine de la révolution industrielle, l'énergie bon marché et abondante accélérant la transformation de la matière : pour l'énergie, nous en voyons aujourd'hui l'importance et la fragilité avec le changement climatique et la crise.

Un troisième axe, plus récent, est constitué de la connaissance, fondée sur les données et leurs échanges, et la propriété intellectuelle qui s'y rapporte. Comme l'a souligné Anne-Sophie Taillandier, il y a quinze ans deux des dix principales capitalisations boursières tiraient déjà une part essentielle de leur activité de données, avec le développement du cloud ; elles sont huit sur dix aujourd'hui, et disposent des plus grandes masses de capitaux pour structurer de vastes pans de l'économie. De plus avec des biens non rivaux et des coûts marginaux de diffusion presque nuls, sont apparues de gigantesques plateformes. Simultanément, ont surgi des questions de souveraineté, de concurrence, et même de démocratie, liées à ces plateformes, et des demandes de réappropriation de leurs données par les individus comme par toutes sortes d'entreprises. Dès lors apparaissent deux besoins : d'une part mieux évaluer les enjeux ; d'autre part mettre en œuvre des régulations et une gouvernance appropriée.

Un besoin de régulation, donc, mais à quel niveau ? Clairement, le risque existe que les États membres de l'UE légifèrent de manière non coordonnée sur les questions liées aux données, ce qui intensifierait la fragmentation du marché unique, entraînant potentiellement une dépendance accrue de l'Union vis-à-vis des États-Unis et de la Chine. C'est pourquoi la Commission Européenne a préparé un train de mesures sur la stratégie numérique de l'Union : le « digital compass 2030 », imposant de développer en Europe trois piliers simultanément : la puissance de calcul (portée par des efforts de R&D, et par une volonté de réduire la dépendance en semi-conducteurs), la connectivité sécurisée (le cas de l'hôpital de Corbeil vient de rappeler s'il en était besoin son importance), et le contrôle de nos données.

C'est de ce troisième volet qu'il a été question ce matin. Il était déjà porté par la directive sur les données ouvertes du 20 juin 2019, qui concerne exclusivement la réutilisation des données du secteur public librement communicables, mais il fallait aussi traiter des données détenues par des organismes du secteur public qui sont soumises à des droits d'autrui et ne relèvent donc pas du champ d'application de cette directive¹. (Il s'agit notamment des secrets d'affaires, des données à caractère personnel, des données protégées par des droits de propriété intellectuelle, de ce qui concerne le secret statistique). Deux textes à portée horizontale y concourent : le règlement du parlement européen et du conseil sur la gouvernance européenne des données, ou Data Governance Act, qui a été adopté en mai 2022, et sera applicable en septembre 2023. Ce DGA porte - schématiquement - sur les données publiques protégées (lorsque les acteurs du secteur public décident d'y donner accès), et sur les intermédiaires en matière de partage de données privées, thème que ce colloque a illustré. Le Data Act, quant à lui est en négociation.

¹ <https://www.consilium.europa.eu/fr/press/press-releases/2022/05/16/le-conseil-approuve-l-acte-sur-la-gouvernance-des-donnees/> L'acte sur la gouvernance des données créera un mécanisme permettant de réutiliser en toute sécurité certaines catégories de données du secteur public soumises à des droits d'autrui. Il s'agit notamment des secrets d'affaires, des données à caractère personnel et des données protégées par des droits de propriété intellectuelle. Les organismes du secteur public autorisant ce type de réutilisation devront être dûment équipés, sur le plan technique, afin que le respect de la vie privée et la confidentialité soient pleinement préservés. Ces données n'étaient pas couvertes par la directive de 2019.

Ces deux textes visent à développer un marché unique de la donnée en Europe en soutenant l'accès, le partage et la réutilisation responsables, dans le respect des valeurs de l'Union européenne et notamment la protection des données personnelles.

Le *Data Governance Act* vise plus précisément à favoriser le partage des données, personnelles et non personnelles, en mettant en place des structures d'intermédiation. Ce règlement comporte :

- un encadrement ainsi qu'une assistance technique et juridique facilitant la réutilisation de certaines catégories de données protégées du secteur public (informations commerciales confidentielles, propriété intellectuelle, données personnelles) ;
- une certification obligatoire pour les fournisseurs de services d'intermédiation de données ;
- une certification facultative pour les organismes pratiquant l'altruisme en matière de données.

Le règlement sur les données (*Data Act*), dont la proposition par la Commission date de février 2022. Son agenda actuel pourrait conduire à une adoption sous présidence espagnole, et une entrée en vigueur fin 2024 ou début 2025. Il vise, comme nous l'avons vu au cours des exposés, à **assurer une meilleure répartition de la valeur issue de l'utilisation des données personnelles et non personnelles entre les acteurs de l'économie de la donnée, notamment liées à l'utilisation des objets connectés et au développement de l'Internet des objets.**

Plus précisément le projet entend traiter de quatre sujets :

- **Faciliter le partage de données privées entre entreprises (B2B) et avec les consommateurs (B2C)**, en fixant notamment une obligation de rendre accessibles les données générées par l'utilisation des objets connectés et services connexes, en contrepartie d'une compensation juste et équitable.
- **Permettre l'utilisation des données détenues par les entreprises et, sous réserve de justifier d'un besoin exceptionnel, par les organismes publics des États membres et les institutions, agences ou organes de l'Union.**
- **Faciliter le changement de fournisseur de services de traitement de données (*cloud* et *edge computing*)** par l'encadrement des relations contractuelles entre les fournisseurs de services et les consommateurs, et notamment par la suppression progressive des frais liés au changement pour le consommateur. Le projet prévoit **de normes d'interopérabilité** pour les données et leurs réutilisations entre les secteurs, condition nécessaire au partage effectif des données.
- Enfin, **mettre en place des garanties contre les accès illicites de gouvernements de pays tiers aux données non-personnelles contenues dans le cloud.** Des avancées ont eu lieu en juillet et septembre tendant à traiter la question des rapports asymétriques entre PME et grands détenteurs de données.

Les approches horizontales comme le *Data Governance Act* et le *Data Act* montrent l'avantage du partage de données, et Laurent Lafaye nous l'a illustré dans l'exemple de Dawex, ou Anne-Sophie Taillandier et Alban Schmutz s'agissant du cloud et de l'initiative Gaia-X. Julien Chiaroni nous a aussi expliqué l'importance du partage des données dans la perspective de l'entraînement des outils d'intelligence artificielle.

Ces approches laissent aussi place à des problématiques, et des régulations plus sectorielles. Xavier Vamparys a ainsi évoqué, dans son article des *Annales*, les enjeux de l'assurance et l'évolution que l'IA apporte à ce métier, en faveur des assureurs, tout en posant des limites éthiques. Jordan Cartier nous a décrit le cas des transports. Comment ceci s'interface avec d'une part des plateformes plus vastes de services, et d'autre part l'autorité de régulation des transports, à la recherche d'un équilibre entre plus de concurrence et d'interopérabilité en faveur des consommateurs. Le besoin de simplicité s'accommode mal en pratique d'un trop grand nombre d'opérateurs, et le coût d'infrastructures informationnelles peut être élevé pour disposer d'informations pertinentes en temps réel. Renaud Font en écho à l'article de Sébastien

Picardat dans les Annales des mines, a décrit des enjeux de la libre circulation des données agricoles. La santé, l'énergie, impliquent aussi que leurs spécificités soient prises en compte, comme leurs enjeux éthiques et géopolitiques. Le colloque des académies des technologies, CAETS 2022 qui vient de se tenir à Versailles sur les technologies de rupture dans les technologies pour la santé l'a clairement illustré dans le cas de la santé, en insistant sur l'importance du partage des données.

Nous voyons bien que les enjeux économiques, sociétaux, de souveraineté, sont considérables, et que nous sommes dans une période charnière. L'article de Jean-Philippe Mochon dans les Annales montre aussi que les interprétations du droit sont encore en cours de stabilisation, avec des divergences qui peuvent demeurer entre Etats dans la transcription de directive, par exemple dans le cas des droits d'auteur et droits voisins. Naturellement, il importera de s'assurer que les dispositifs mis en œuvre ne peuvent être contournés en dehors de l'Union. Il y a donc un vaste champ de recherche qui reste ouvert et pour lequel les initiatives de la chaire de l'Université Paris-Dauphine qui nous accueille aujourd'hui sont les très bienvenues ;

Je remercie très chaleureusement Eric Brousseau, et Jean-Yves Ollier pour avoir rendu cet évènement possible. Le conseil général de l'économie pour sa part s'attache à contribuer à ces débats et à la coconstruction de régulations au service de notre économie et de notre société, notamment de façon ouverte via les Annales des mines que dirige maintenant Grégoire Postel-Vinay. Merci à tous.



Chaire Gouvernance et Régulation
Fondation Paris-Dauphine
Place du Maréchal de Lattre de Tassigny - 75016 Paris (France)
<http://chairgovreg.fondation-dauphine.fr>