

Les 11 propositions du GFII dans le cadre de la mission Bothorel sur la politique de la donnée et des codes sources

Point 1 - Données et souveraineté : toutes les données ne se valent pas

Les données sont des actifs numériques essentiels, tant pour les Etats que les entreprises ou les collectivités territoriales. Même si les données publiques sont dites « de libre parcours », il n'en demeure pas moins que **certaines, dans des domaines stratégiques, ne peuvent être communiquées de manière identique et aveugle** pour toute réutilisation, pas forcément bienveillante. **La politique « open data »** actuelle, basée sur une liberté totale et sur l'absence de contrôle **n'est pas viable** à terme pour ce type de données.

Le GFII attire l'attention sur le fait qu'il existe une concurrence internationale pour la fourniture de données, par exemple, sur les données géographiques et météorologiques, ce qui renforce d'autant plus la nécessité de maintenir un haut niveau de qualité pour assurer l'attractivité des acteurs français sur ces marchés, et garantir (aussi !) l'exercice souverain des politiques publiques qui s'appuient sur leurs données.

Point 2 - Données et potentiel économique : toutes les données ne se valent pas.

Toutes les données n'ont pas un potentiel économique identique : rien de commun entre les horaires des piscines, la localisation des places de parking et le fichier SIRENE. Intrinsèquement les données recèlent des potentialités différentes. Or, les entreprises, qu'il s'agisse de start-up ou d'entreprises existant de longue date, **ne peuvent développer des projets et de nouveaux services basés sur la data sans que la pérennité de la fourniture de donnée ne soit confirmée par le producteur et que les qualités essentielles de la donnée soient maintenues, voire améliorées.**

Point 3 - Exploiter des données : le besoin d'une maîtrise technique du domaine/ sujet traité

Le terme « open data » laisse souvent penser qu'utiliser des données est simple. Or, outre les aspects « traitement » de données, **certains jeux de données nécessitent une maîtrise technique de la « matière »**, a fortiori si plusieurs sont croisés entre eux ou utilisés pour l'IA, qui exige une professionnalisation de l'exploitation et de la connaissance de la donnée.

C'est la raison pour laquelle l'on retrouve dans la chaîne allant du producteur de données à l'utilisateur final de plus en plus de « spécialistes intermédiaires », à même de traiter des données en masse, de les normaliser, de les désambigüiser, de les hiérarchiser, de leur apporter une valeur ajoutée, manuellement ou automatiquement, bref « à les faire parler ».

Point 4 - Les besoins des acteurs économiques sont distincts de ceux du grand public et requièrent une qualité des données supérieure

Dans ce cadre professionnel, la « data » devient vitale, car elle permet à l'entreprise ou à l'administration, une meilleure connaissance de son environnement, de ses concurrents, de ses clients/usagers et, surtout, permet le développement de nouveaux services à valeur ajoutée, gages d'un développement numérique. Pour y parvenir, il est capital que la donnée soit de qualité. Comment définir la qualité ? C'est une donnée fiable, tenue à jour, exhaustive et dont le producteur garantit la permanence de la qualité, la maintenance et la pérennité de sa fourniture.

Sans qualité, pas de développements d'une IA fiable (risque de biais), pas de nouveaux outils d'aide à la décision.

Point 5 - Il existe plusieurs types de réutilisateurs dont les environnements économiques et juridiques sont différents, ce qui justifie une mise à disposition des données différenciée

La réutilisation des données ne relève pas d'un bloc monolithique et il convient de prendre en compte les différents cercles susceptibles d'exploiter des données pour des finalités/usages différents. **Ainsi le GFII travaille-t-il plus particulièrement sur l'économie et l'exploitation de la donnée par des professionnels, publics ou privés, pour des professionnels et dans une optique de production de services ou d'applications professionnels**, avec le niveau d'exigence afférent, d'ailleurs décrites dans l'European Data strategy définie par la Commission européenne en février 2020 : administration à administration (GtoG), administration à entreprise (données publiques G to B), entreprise à entreprise (B to B), entreprises vers administrations/données d'intérêt général (B to G). Il est essentiel que cela soit reconnu à tous points de vue, car les usages (et les contraintes juridiques et d'anonymisation) sont forcément spécifiques.

Point 6 - Repensons le modèle économique des données publiques !

La mise à disposition gratuite de données est bien sûr séduisante mais peut s'avérer contre-productive quand il s'agit, sur le long terme, de produire et diffuser des données de qualité, interopérables et via des modes de diffusion (dont les API) susceptibles de pouvoir gérer un grand nombre de requêtes.

De plus, les difficultés de financement ne doivent en aucun cas générer une baisse de qualité dans la mise à disposition des données car ceci impacte tous les produits et services développés par les réutilisateurs, ainsi que leurs clients finaux. Dans le cadre de l'ouverture des données publiques au sens large (ie pas seulement les données de référence / jeux de données de forte valeur), il faut que l'Etat évalue l'impact des investissements nécessaires et n'hésite pas, s'il ne peut les faire supporter au budget général, à nouer des partenariats public-privés (délégation de service public) ou autorise les acteurs publics concernés à :

- **demander des redevances dans le cas où les données sont réutilisées à des fins commerciales** dès lors que la base de calcul et le montant des redevances sont publics et que leur évolution est évaluée par une autorité indépendante sur la base de critères précis, objectifs, transparents et vérifiables et que leur gestion est simple à mettre en œuvre pour les deux parties

et/ou

- **développer des services optionnels et payants associés à la mise à disposition de données publiques** (format, niveau d'anonymisation, fréquence de délivrance...)

Point 7 - Articuler protection des données personnelles et données publiques - Mettre fin aux incohérences normatives

Il y a des incohérences de textes, tant au niveau français qu'europpéen, sur l'anonymisation ou non, et l'accès et la réutilisation de données personnelles mises à disposition dans des jeux de données publiques accessibles en open data.

Et ce, tout au long de la chaîne de traitement, soit du producteur de données publiques à l'utilisateur final de la donnée.

Il y a, par exemple, un besoin de clarification normative en ce qui concerne les raisons sociales d'entreprises qui correspondent au nom de leur dirigeant (entreprises individuelles) et, partant, sur la possibilité de réutiliser ces données qui perdraient du coup leur caractère de données personnelles à partir du moment où **leur utilisation est en lien direct avec l'activité de l'entreprise.**

Ce même besoin de clarification entre données personnelles « privées » / données personnelles accessibles au public / données personnelles de publicité légale destinées à l'information des tiers / données personnelles non anonymisées mises à disposition dans des jeux de données publiques disponibles en open data (art D 312-1-3 du CRPA) touche de nombreux domaines : les coordonnées d'un maire, les informations personnelles dans les jeux de données relatifs aux entreprises, dans une décision de jurisprudence etc.

Il est donc nécessaire que les interactions entre les textes puissent être prises en considération dès la rédaction des projets.

Point 8 - Investir dans les données susceptibles d'être les plus réutilisées en partant de cas d'usage réels (« use case »)

Selon la « stratégie européenne pour la donnée », « la valeur des données tient à leur utilisation et réutilisation ». Or, si les données sont au cœur du numérique, toutes ne se valent pas.

C'est la raison pour laquelle **le GFII considère qu'il est essentiel de réfléchir sur la base de cas d'usage, bien réels, qui permettent de faire remonter les besoins et contraintes des ré-utilisateurs.** Une réglementation unique pour tous les secteurs ne peut pas poser les conditions d'une qualité optimale de la donnée.

Ces cas d'usage ne peuvent être étudiés que dans le cadre d'un dialogue entre le(s) producteurs(s) de données, public(s) ou privé(s), et les réutilisateurs, sans préjugés de part et d'autre, dès lors que « *chaque domaine a ses propres spécificités* » (Stratégie européenne de la donnée). Il nous semble ainsi plus judicieux de travailler d'abord sur la base d'approches verticales, par secteur, qui doivent pouvoir intégrer des acteurs qui pourraient ne pas en être, à première vue, sans se priver de s'inspirer de ce qui aura bien fonctionné pour tel ou tel secteur pour en tirer des principes communs. Et ensuite, de travailler sur l'approche horizontale.

Point 9 - Envisageons des droits d'accès aux données adaptés en fonction de la finalité de réutilisation

Un droit unique, identique pour tous, en arrive de facto à un droit restreint pour une partie des cas d'usage, ce qui apparaît contre-productif par rapport à l'objectif visé de développer l'économie de la donnée. **Cela ne devrait pas empêcher que ces droits et devoirs soient équitables, pratiques et clairs**, ni même que les données puissent circuler à l'intérieur de l'UE et entre les secteurs. **Ainsi, il serait pertinent de prévoir différents niveaux d'accès aux données, de droits et devoirs en fonction des finalités de réutilisation et des cercles de réutilisateurs (voir point 5).**

La création d'espaces européens communs des données par secteur va dans le bon sens. Il conviendra toutefois de ne pas avoir une interprétation trop stricte des dits secteurs en termes d'entités à inclure. Il conviendra toutefois de ne pas avoir une interprétation trop stricte des dits secteurs en termes d'entités à inclure. Cela nous semble d'autant plus important que les données sont des enjeux de souveraineté et de puissance dont la communication à des tiers non identifiés peut s'avérer dangereuse.

Point 10 - Prévoir une mise à disposition des données d'intérêt général reposant sur le consentement des entreprises et leur indemnisation

La crise du Covid19 a clairement montré l'intérêt pour l'action publique au sens large de pouvoir s'appuyer sur des données de source privée ou données d'intérêt général. **Or, les données des sources privées représentent pour leur producteur un patrimoine numérique essentiel et un atout économique indispensable.**

Là encore, le GFII encourage à travailler sur la base de cas d'usage et surtout pas sur la base de grands principes généraux. C'est de la multiplicité des cas d'usage et des retours d'expérience que des principes généraux communs pourront éventuellement être tirés.

Le sujet de l'indemnisation devra être exploré, d'autant qu'il faudra certainement faire des développements pour répondre au besoin de l'Etat au sens large. Exactement comme les acteurs publics doivent faire des développements pour permettre une réutilisation efficace de leurs données.

Rendre le BtoG obligatoire serait contre-productif ; il conviendra donc d'être incitatif et non de procéder de manière coercitive. Ainsi, la mise à disposition de données des données d'intérêt général ne pourra à priori fonctionner que sur la base du volontariat des acteurs économiques privés et leur indemnisation

Point 11 - Rendre le contenu des licences cohérent, clair et transparent

Le GFII alerte par ailleurs sur la teneur de certaines licences, même publiques, et autres conditions de réutilisation. **Les données ne seront largement réutilisées que si les « règles du jeu » sont claires**, précises et transparentes, accessibles même sans disposer de compétences avancées en droit.

***A propos du GFII:** Association Loi 1901, le GFII a pour objectif le développement de l'économie et l'emploi en France grâce au numérique. Il œuvre pour que les besoins/attentes des professionnels de contenus et de données dans un environnement B to B soient pris en compte par les pouvoirs publics.*

Pour en savoir plus: <https://www.gfii.fr/fr/>, ou suivez le GFII sur [LinkedIn!](#)

Contact : Anne-Claire Marquet, Déléguée générale, dg@gfii.fr ou 06 99 29 51 87