

Actes de la journée du Réseau d'Observation Sociale et Sanitaire Bourgogne Franche-Comté

Données sanitaires et sociales :
comment faire face à la multiplication des possibles et des usages ?

13 octobre 2022

Préambule

Les données participent à la construction des politiques publiques en général, et a fortiori des politiques sociales et sanitaires. C'est pourquoi des services d'études et de recherches existent au sein des ministères au niveau national comme au niveau régional.

Des évolutions sont notoires.

- La manière dont les données sont produites, traitées et analysées pour mettre en œuvre les politiques publiques est en train d'évoluer fortement
- Les attentes de la part de la société pour accéder aux données, dans une logique de transparence, s'accompagnent également d'un renforcement des exigences de confidentialité, et de protection de la vie privée
- Une production accrue de donnée qui doit être valorisée, y compris de manière marchande
- Le progrès technique permet directement aux non spécialistes d'accéder et d'exploiter de la donnée
- Une adaptation du métier de statisticien : Pour répondre à ces transformations, le métier de statisticien/chargé d'études... (apparition du métier de data scientist) évolue en développant de nouvelles compétences (programmation, pédagogie, communication,) et aussi peut être plus fondamentalement dans son positionnement (peut-être de moins en moins producteur de donnée et de plus en plus analyste, interprète, "oracle", communiquant, etc.)
- Des évolutions qui impactent même les services métiers : Ces transformations induites par le numérique impactent la statistique mais également les services métiers (confidentialité, sécurisation des données, outillage numérique de plus en plus perfectionnés...) qui ont des accès facilités à des outils statistiques, cartographiques..., et qui sont de plus en plus producteurs de data (+ même que les services statistiques). cela peut mettre en difficulté les services mais également ouvrir des opportunités.

Cette journée aborde ces transformations à travers des retours d'expériences d'utilisateurs multiples/divers (services statistiques de niveau régional et national, le monde de la recherche...)

Introduction



Philippe BAYOT, directeur régional délégué de la Directions régionales de l'économie, de l'emploi, du travail et des solidarités (Dreets) Bourgogne Franche-Comté

Cette journée est ancrée dans les pratiques quotidiennes des services statistiques, leurs évolutions et leurs conséquences. Sont à noter :

- Une tendance des politiques publiques à vouloir « monitorer » l'action publique qui s'intensifie. Celle-ci est utile et répond aux besoins de savoir, par exemple, comment l'action se déroule sur un territoire.

Mais il peut y avoir un problème de temporalité entre l'action et les données, qui ne coïncident pas toujours. Le « politique » a en effet parfois besoin des données en urgence, lesquelles ne sont pas toujours disponibles et l'analyse, elle, nécessite un temps et de la distance par rapport à l'action.

- Le rapport entre production de données et interprétation. Les données/informations parviennent de plus en plus de manière indifférenciée entre les « généralistes » et les « spécialistes ». Il y a donc nécessité pour les seconds (statisticiens) de disposer de compétences et d'une « pédagogie de la donnée » pour que les premiers se saisissent de manière éclairée de leur sens et de leur interprétation. A ce titre, le rôle important des services études est à rappeler.

- Le rapport nécessaire entre production de la donnée, interprétation et innovation. Dans notre monde « post COVID » où les outils de d'information et de communication ont été démultipliés, il y a pour les services d'étude qui doivent pouvoir apporter de l'innovation dans la façon d'animer les politiques publiques, en produisant des outils innovants de transmission, d'information et d'intégration des analyses (capsules, webinaires, etc.). Le champ est large, c'est une occasion à saisir pour poursuivre de manière renouvelée le raccordement entre ce qui relève de l'étude et ce qui relève de l'action.



Cédric DUBOUDIN, Directeur de la direction de l'innovation et de la stratégie de l'Agence régionale de santé (ARS)-Bourgogne Franche-Comté

La place de la donnée dans la décision publique, « data to action », est devenue centrale et prégnante. La crise du Covid a conduit à davantage de monitoring par la donnée et le besoin de s'appuyer sur des data et des outils de simulation. Il s'agit de fonder la décision sur la rationalité de la donnée plus que sur le ressenti et le subjectif. L'utilisation de la donnée pour la décision publique reste relativement faible en santé, comparée à d'autres domaines privés et publics. Cette nouvelle donne amène les interrogations suivantes :

- La relation entre data scientist et décideurs et personnel « métier »

- Une interrogation sur les méthodes ; la donnée brute aide rarement le décideur mais l'analyse permet de montrer des corrélations, des zones de besoins, d'anticiper et de simuler des évolutions. Il est nécessaire de s'entourer de compétences en ingénieur. De plus, la frontière est parfois ténue entre données étudiées à des fins de politiques publiques et celles à des fins de recherche.

- La nature de la donnée ; les besoins évoluent entre une donnée fixe et celles interopérées avec d'autres pour accéder à différentes dimensions autour de l'individu.

- Le cycle de la donnée ; celui de sa production, celui de son suivi à travers un outil, celui de l'interrogation du résultat.

L'accès à des données fines, riches et pérennes permet, par leur traitement, de fonder les décisions sur des éléments rationnels.

Le Health Data Hub : une plateforme de données de santé pour la recherche d'intérêt public



Nicolas Bellière, chef de cabinet de la direction générale, Health Data Hub

Les porteurs de bases de données ont un réflexe propriétaire de leur donnée. Or les données de santé doivent d'abord servir au patient et à la qualité des soins. Dans cette optique, le Health Data Hub (HDH) représente un guichet unique d'accompagnement et d'accès à des données notamment du système national des données de santé (SNDS). Il met à disposition un catalogue de données, une plateforme, des outils, et soutient le déploiement d'une culture commune entre acteurs utilisateurs de données. Il intervient en amont, en vérifiant la complétude du dossier pour passage au Comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé) (Cesrees) et, en aval, à l'occasion par exemple d'un accès au SNDS ou de la constitution de base de données appariées à celles du SNDS qu'il constitue lui-même. Une fois ces bases appariées elles sont mises à disposition dans un espace sécurisé créé pour l'utilisateur.

Le HDH est donc une plateforme sécurisée sur laquelle on trouve :

- Un catalogue de données avec des listes de bases de données
- Des espaces projets qui sont des « bulles sécurisées » notamment équipées de logiciels de traitement et d'analyse mis à jour annuellement pour répondre aux besoins des utilisateurs
- un pré-traitement des données (en particulier dans le cas des données appariées pour lesquelles des variables ont été retenues pour l'appariement).

Le HDH ne travaille pas sur la question statistique. Il se présente comme un tiers digne de confiance et facilitateur de l'écosystème qui permet de faire avancer la recherche au niveau médical. Il bénéficie de financement public et privé. Ces derniers sont mobilisés dans le cadre de valorisation particulière de base de données (appariement notamment).

Il permet 3 grands types d'analyse : épidémiologiques, socio-économiques, méthodologiques.

Il compte actuellement 3 programmes de travail :*

1. Appel d'offre à destination des établissements de santé en vue d'améliorer les prescriptions médicamenteuses
2. Un appel à projet pour le déploiement d'entrepôts de données de santé (en lien avec le ministère de la santé)
3. La création d'un entrepôt de données de la médecine de ville

Tous les projets mobilisant des données du HDH font l'objet d'une fiche de déclaration accessible et consultable sur le site internet qui en compte actuellement 6000 (fiches).

Son équipe compte 83 personnes dont 25 travaillent sur les bases de données. Les profils de l'équipe sont variés et composés pour moitié de profils de data manager/technicien data

Le HDH travaille également sur la création d'un espace européen des données de santé et la question des conditions du transfert de données d'un pays à un autre avec l'idée d'un formulaire commun d'accès aux données. Un autre axe de travail concerne l'implication des citoyens et ses modalités (étudiants, journalistes, think tank), avec la mise à disposition de formation sur les données existantes.

L'évolution du métier de statisticien du champ social face aux transformations numériques



Olivier LÉON, adjoint au sous-directeur de l'observation de la solidarité, Drees

La Direction de la recherche, des études, de l'évaluation et des statistiques (Drees) est une direction de l'administration centrale des ministères sanitaires et sociaux. Elle agit sous la tutelle des ministères de la Santé et de la Prévention, des Solidarités, de l'Autonomie et des Personnes handicapées, du Travail, du Plein emploi et de l'Insertion, de l'Économie, des Finances et de la Souveraineté industrielle et numérique (pour l'exercice de ses attributions relatives aux finances sociales). Elle produit des statistiques, réalise des synthèses et études, contribue à l'évaluation des politiques publiques.

La multiplication des bases de données dans le champ social, répondant aux besoins des usagers, des établissements, des administrations et des financeurs, permet d'accroître le champ des possibilités d'analyse comme faire des simulations sur l'accès aux droits, améliorer le traitement de dossier quand il requiert plusieurs intervenants, cartographier des services ou des structures.

Le champ compte des systèmes construits en silo, mobilisant des micros données, au sens de données fines, sur le sujet traité. Dans les systèmes d'informations la temporalité est peu présente. Différentes questions sont posées :

Les données de nature différente, et les singularités des données statistiques et celles de gestion. La donnée de gestion n'est pas forcément traitable « en l'état », pour être diffusable. Elle peut être construite sur la base de concepts différents, compter des oublis, nécessiter des redressements. Elle repose sur un travail préalable pour permettre l'agrégation des unités, ce qui mobilise un temps souvent incompris par les commanditaires. Le statisticien peut réunir des données pour effectuer des évaluations à long terme (de dispositifs par exemple) et la variable temps peut, là aussi, avoir un rôle important, du fait des évolutions sociétales existantes à court, moyen et long terme.

Le statisticien peut réaliser des analyses transversales et non en silo. Il raccroche les données à des caractéristiques plus larges des individus, de leurs territoires de vie. Il a donc besoin d'une clé d'appariement alors que le concepteur d'un système de données de gestion aura lui, anonymisé son système, au regard des contraintes du RGPD. C'est en amont, lorsque le système d'informations est conçu que le statisticien doit définir et faire valoir ses besoins et ses droits d'accès. Les possibilités d'appariement, outre l'analyse plus riche, permettent de disposer d'informations collectées une fois et non pas à de multiples reprises. Le statisticien doit se muer en ingénieur des systèmes d'information pour optimiser et rationaliser les demandes d'informations auprès des différents opérateurs. Il doit être à l'affût de nouveaux systèmes qui apparaissent, aider à définir les concepts et le champ des données, documenter les données et les dispositifs de traitement (collecte, analyse, redressement, corrections des données).

La question du temps pour répondre au calendrier des demandes est mise en perspective avec le risque d'erreur : quand on est trop rapide, on peut rendre des résultats erronés qui conduisent à des décisions basées sur de faux constats. De même l'évaluation nécessite du temps de traitement des données.

Le droit d'accès du statisticien public à des données identifiantes, appariées à d'autres systèmes est intéressant pour être économe lors de questionnaire d'enquête. Cependant faire valoir ce droit d'accès aux données n'est pas toujours simple pour le statisticien, a fortiori depuis les dispositions du RGPD.

Une plateforme de support dans l'analyse statistique dans un institut en sciences sociales : fondamentaux et évolutions



Arnaud BRINGÉ, responsable du service des méthodes statistiques, Ined

L'Ined est un organisme de recherche publique en sciences sociales. Il est organisé avec des unités de recherche thématiques en problématiques démographiques (fécondité, mortalité, histoire, économie ...), et des services d'appui spécialisés proposant une offre de services transverses aux différentes unités de recherche. Le service Méthodes Statistiques s'inscrit depuis 2007 dans cette logique offrant des services collectifs (formations internes, séminaires, sites Web dédiés ...). Ces dernières années ont été marquées par la prise en charge d'outils collaboratifs, et la volonté d'une politique de transparence dans le cadre de la science ouverte.

Pour faire face à une dualité entre expert métier et chargé d'études aux disciplines et statuts variés (chercheur, étudiant, statisticien, démographe), et maîtriser l'accélération des possibles et des usages, l'Ined a créé une plateforme de supports pour les utilisateurs de données, service d'appui composé d'ingénieurs en science des données apportant un support en méthodes et outils statistiques en sciences sociales. Elle travaille sur une pluralité de données ; d'enquête, administratives, du SNDS, d'entretiens, des données de réseau, parfois scrapées¹ sur le web. Cette plateforme constitue un service d'appui à la recherche et travaille, au-delà de l'Ined, en partenariat avec des chercheurs associés (Inserm, Insee, EHESS, etc.). Avec chacun, elle vise à développer, à l'occasion des projets, des offres spécifiques (statistiques de comparaisons de données internationales, méthodes d'analyses, typologies, etc.). Elle répond à des besoins statistiques différents et liés aux disciplines présentes, et mobilisent également des méthodes nombreuses (analyses biographique et multiniveaux, petits ou grands échantillons, modèles statistiques appliqués, recherches de liens de causalité, analyse de séquence, analyses de réseaux, typologies ...). Face aux données de plus en plus massives, leur représentation de manière simple est une préoccupation forte. En partenariat avec des diffuseurs de données tels le Datalab de l'Ined, ou les partenaires² des PUDs (Plateformes Universitaires de Données) et PROGEDO³. La datavisualisation occupe une place de plus en plus importante avec des outils tels que Leaflet⁴ ou mapsf⁵. L'Ined utilise les outils R, MAGRIT⁶ et QGIS pour les représentations cartographiques.

L'Ined s'inscrit dans une démarche de traçabilité des données pour garantir la fiabilité, la compréhension et l'analyse. De nouvelles méthodes telles que le scraping visant à détecter et utiliser les données du Web (même si ceci nécessite des questions juridiques et informatiques complexes), et des outils accessibles en ligne tels que R et ses dérivés, constituent des opportunités de développement et de partenariats.

¹ Le « web scraping » est une technique d'extraction automatique des données d'un ou plusieurs sites web. Elle peut être effectuée à l'aide d'un logiciel, d'un outil ou d'une application, mais elle peut aussi être réalisée par des développeurs. Elle permet d'extraire des datas en grande quantité de façon structurée. Cette pratique est notamment encadrée par des directives de la Cnil réunies dans le document « La réutilisation des données publiquement accessibles en ligne à des fins de démarchage commercial », 30 avril 2020.

² CNRS, Ined, au sein du Consortium of European Social Science Data Archives (CESSDA)

³ Catalogue de données en sciences humaines et sociales pour la recherche qui met à disposition plus de 1500 jeux de données.

⁴ <https://data.progedo.fr/>

⁵ <https://leafletjs.com/>

⁶ <https://riatelab.github.io/mapsf/>

⁶ <https://magrit.cnrs.fr/>

La donnée fait-elle ou non l'évaluation ?



Direction régionale
de l'économie, de l'emploi,
du travail et des solidarités

Tony FOGLIA, chargé d'études, docteur en économie de la santé, ORS BFC

Lionel DURAND, chef du service Études Statistiques Évaluation, Dreets BFC

Les données sont indispensables à l'évaluation de l'action publique. L'amélioration de l'accessibilité de l'information (statistique en particulier) peut donner le sentiment que l'évaluation sera simplifiée à l'extrême, voire qu'on pourrait aller vers une évaluation automatique. Sans préjuger de ce qu'il pourra advenir, il convient toutefois de rappeler que la donnée n'apporte rien à l'évaluation si elle ne s'inscrit pas dans un cadre méthodologique qui articule des dimensions scientifiques, techniques et politiques.

En 2015, la société française de l'évaluation soulignait le défi que représentait le « big data » face à l'évaluation des politiques publiques⁷. Avec d'un côté un accès plus grand à des données de plus en plus vastes et nombreuses, et de l'autre, des différences sur les données, méthodes, concepts, qualité d'information. De plus, l'évaluation cible des processus, résultats et plus globalement un impact des politiques publiques, pas toujours simples à mesurer, notamment du fait de la temporalité différente entre la mise en œuvre d'un programme, d'une action et celle de ses effets.

Ainsi, l'évaluation est sensible au calendrier, mais contrainte par le délai d'accès aux données, lui-même dépendant des systèmes de données disponibles et/ou construits pour l'action. Les systèmes des décideurs publics permettent parfois de disposer d'indicateurs utiles à l'évaluation mais pas toujours. Dans ce cas, il faut construire des systèmes de données et de collecte complémentaires. Cette démarche suppose des compétences, du temps, des moyens ad hoc. La question de la donnée, des indicateurs, est à mettre en perspective avec la gouvernance politique, les méthodes de sa planification et de son évaluation. Alain Supiot, dans son ouvrage « La gouvernance par les nombres »⁸ analyse notamment le poids de ces mesures dans la comparaison avec une norme idéale qui encadrerait la réalisation d'objectifs mesurables.

Or au-delà des indicateurs et objectifs mesurés, l'intérêt de l'évaluation repose sur l'analyse réflexive sur l'action et sa mise en œuvre. Dans ce sens la complémentarité des méthodes, quantitatives et qualitatives et leur intérêt pour l'évaluation sont à souligner. La combinaison d'une approche statistique et sociologique fait écho à un autre constat : le Big data permet d'accéder à un format de généralités qui doivent être complétées par des approches analytiques des pratiques pour aboutir à une information plus complète et donc une évaluation plus juste et précise de l'action. Pour l'évaluation, un service statistique n'est pas seulement présent pour fournir un indicateur ; il peut intervenir pour rendre un dispositif évaluable, c'est-à-dire expertiser les systèmes d'informations et dessiner les orientations d'un système ad hoc.

Selon les domaines, l'évaluation bénéficie parfois d'une certaine antériorité. Par exemple nombre d'actions du domaine de la prévention en santé ont déjà été évaluées. Les questions nécessitant des évaluations complémentaires portent plus sur les conditions positives de la transférabilité des actions ou programmes que sur leurs effets déjà connus par ailleurs. Pôle emploi signale sa situation entre un cahier des charges national de l'évaluation et sa réalisation régionale, tout comme celle d'évaluateur interne qui convoque la position de « juge et parti », qui conduit à confier certains travaux d'évaluation à des intervenants externes.

⁷ Le défi de la révolution foudroyante du numérique Big Data et Open Data, quels enjeux pour l'évaluation ? http://www.sfe-asso.fr/sites/default/files/document/file/data_synthese_om-v3.pdf

⁸ Paris : Fayard, 2015, 512 p.

Le statisticien face aux enjeux posés par l'utilisation de sources administratives



Direction régionale
de l'économie, de l'emploi,
du travail et des solidarités



Émilie VIVAS, adjointe au chef du service Études Statistiques et Évaluation, Dreets BFC

Brice NAFETAT, chargé d'étude, Pôle Emploi

Dans l'environnement du statisticien, les sources administratives occupent une place croissante. Elles permettent de compléter voire de se substituer aux données d'enquête, elles peuvent également constituer une source nouvelle. Dans un contexte d'explosion des possibilités en matière de données, le statisticien doit faire face à de multiples enjeux qui engendrent une nouvelle façon de travailler.

Les données administratives sont toutes celles produites par l'État. Ce sont des données très centralisées, issues notamment de processus de gestion interne. Pour être utilisables, elles doivent la plupart du temps être « retraitées ». Elles existent généralement pour suivre une action. En même temps que les dispositifs créés par l'État, de nouveaux champs de données peuvent également apparaître. L'Open data permet d'accéder aux données sans attendre qu'elles soient traitées par l'organisme qui les produit. Mais cette possibilité engendre des contraintes de stockage et la nécessité de retraiter soi-même des données brutes. Cet accès plus large à des données multiplie le nombre d'acteurs disposant des mêmes données que chacun va traiter parfois différemment ce qui produit des chiffres parfois différents.

Chacun construit son système administratif et le renseigne à sa convenance, sans cadre commun. Pour traiter ces données il faut les figer à un instant T. Là aussi chacun peut choisir d'arrêter le système à la date qui lui convient.

Face à ces possibilités et ces usages, il y a un environnement à connaître pour utiliser la donnée :

- Quelles sont les règles de gestion à définir, à écrire, pour expliquer ce qui la construit ? Par exemple, certains changements réglementaires peuvent avoir un impact fort sur l'évolution des variables indépendamment du contexte conjoncturel ou entraîner des ruptures de séries de données.
- Quel est son type ? Les notions de stocks et de flux sont à prendre en compte pour leur interprétation, de même que l'existence de données brutes nécessitant des procédures de corrections de variations saisonnières pour apprécier leur évolution.
- Quel découpage géographique est utilisé ? Le zonage administratif peut être différent selon l'administration qui collecte la donnée.

Face aux données administratives, un chiffre reste une estimation d'un événement ou de caractéristiques d'individus, obtenu à partir d'une saisie manuelle ou automatique dans un système d'information. Il faut donc parfois avoir du recul par rapport à sa précision.

Il y a un dialogue à mener entre le statisticien et la personne qui saisit les données pour accéder au sens des informations et veiller à la conformité de la saisie avec l'information d'origine. Ce dialogue est important car l'évolution des systèmes d'information ne suit pas forcément l'évolution des dispositifs, et le sens des données enregistrées peut donc en pâtir.

Le statisticien devrait être associé en amont des évolutions mais leur anticipation reste difficile et combine également l'évolution des outils techniques.

Par exemple, en 2010, la Délégation interministérielle à l'hébergement et à l'accès au logement (DIHAL) a été créée. Cette création a entraîné la modification du système d'information et de gestion de l'hébergement d'urgence (ex 115). Les données ont été enregistrées différemment pour être en phase avec l'ensemble des

missions et dispositifs de la Dihal et non plus strictement sur celles de l'hébergement d'urgence. Cet exemple souligne une nouvelle fois la relation entre le système d'informations et le statisticien.

Développement des besoins en matière de diffusion de données, utilisation de nouveaux outils : les transformations du métier statistique favorisent-elles la coopération ?



Jérôme MATHIAS, attaché statisticien, Insee

Didier CAREL, chef du Département études et statistiques , ARS BFC

De plus en plus utilisée, la plateforme de développement statistique R voit sa communauté d'utilisateurs s'agrandir. Cette solution de développement est modulaire et les nombreux packages et différentes approches de programmation existants peuvent induire des besoins d'échanges, d'appui, de partage de la connaissance accrues. Le package Shiny de R permet le déploiement d'application web pour la diffusion d'informations ou de traitements statistiques. Ce déploiement s'appuie sur des infrastructures dont la mise en place présente des prérequis techniques, des besoins en administration qui peut encourager à les partager.

L'expérience de l'ARS face à l'utilisation de l'outil R est celle d'un développement en équipe, avec la création d'une documentation du ou des codes et des outils de partage de ces codes. R s'appuie sur différents packages face auxquels il faut être attentif, en cas d'utilisation en équipe, à l'installation des mêmes packages sur les différents postes informatiques afin de pouvoir utiliser les mêmes codes. Un environnement de développement commun avec homogénéité des versions peut faciliter le partage du code puis sa publication.

L'accompagnement des utilisateurs de R est important, avec le besoin d'une acculturation commune face aux programmes, outils. Du côté de l'utilisateur des vues R shiny par exemple, la phase d'interrogation des besoins n'est pas à négliger afin d'apporter des réponses satisfaisantes en présentant des données utiles et compréhensibles.

R Shiny est un package gratuit et il est possible de développer des interfaces et de les tester localement sur son poste de développement avec R Studio. Mais pour publier ces interfaces sur internet il faut un serveur équipé de Shiny server. La maintenance d'un tel serveur nécessite des compétences d'administration. L'ARS l'a installée au Grades BFC qui est également hébergeur de données de santé. Une alternative plus simple est l'utilisation d'un compte Shappy.io qui permet d'héberger une application Shiny sans disposer de son propre serveur et avec plusieurs formules d'abonnement possibles dont certaines gratuites.

Au sein du service statistique publique, la DREAL et l'Insee co-animent un club R ouvert depuis juillet 2022 et comptabilisant une centaine de membres. Il permet de partager autour des packages (18700 à ce jour disponibles sur le CRAN) souvent accompagnés d'une (trop) courte documentation. Différents outils de formation et de mutualisation⁹ sont à disposition des personnels de l'administration ; IGOR qui propose une autoformation, le datalab Onyxia¹⁰ qui propose des ressources informatiques, des zones de stockage , de test, des webinaires. De son côté, Utilit R¹¹ fournit une documentation rédigée par des personnels Insee dans une visée collaborative, il est ouvert à tous.

Les économistes de l'université qui utilisent R pourrait contribuer au datalab car les adresses en .cnrs.fr sont autorisées. Pour les adresses de l'université il faut tester leur éligibilité et si elles ne sont pas acceptées faire

9 https://osmose.numerique.gouv.fr/jcms/p_3168022/fr/mte-bourgogne-franche-comte-club-regional-r-interministeriel-en-bfc

10 <https://datalab.sspcloud.fr/home> ; <https://docs.funcamp.sspcloud.fr/>

11 <https://www.utilitr.org/>

une demande au datalab afin de voir si la liste des adresses peut être mise à jour. Le datalab est fait pour du prototypage, il n'existe pas d'assurance sur le niveau de service du datalab. La qualité des interactions informatiques varie en fonction du nombre d'utilisateur présents simultanément. La qualité est meilleure lorsqu'ils sont moins nombreux. Onyxia pourrait être une solution pour l'utilisation de R par des agents de la Dreets qui ne peuvent installer R sur le serveur.

Conclusion et perspectives



Adrien Roussot , chargé d'études, co-pilote du ROSS, ARS BFC

Cette journée a montré combien la statistique était désormais au centre des préoccupations au sein des organismes administratifs. Il faut donc réguler les demandes, appréhender les nouveaux outils, digérer des données de plus en plus importantes.

Il faut donc s'intéresser aux utilisateurs des données dans les organismes mais aussi aux données vulgarisées et restituées auprès du grand public. Celui-ci pourrait mériter un accompagnement à la compréhension de ces données

L'ouverture des données vers la sphère publique est à la fois une tendance et une demande et la question des ressources pour y faire face est tout aussi importante si on veut garantir la qualité de ces données et des analyses.



Direction régionale
de l'économie, de l'emploi,
du travail et des solidarités

**Rina Petrequin, cheffe de projet Observation sociale, Service études statistiques et évaluation (SESE),
co-pilote du Ross, Dreets BFC**

En complément et dans les perspectives qui pourraient découler de cette journée, le domaine de l'intelligence artificielle figure parmi les évolutions à noter et pour laquelle les services statistiques ont besoin d'être accompagnés au changement. En complément, la formation nécessaire à la compréhension et l'interprétation des données restent centrale face à des systèmes de plus en plus complexes.