

# Indicateurs de qualité en matière de production de données : quelques éléments de réflexion

 Pascal Rivière\*

**L**e présent article s'inscrit dans la continuité d'un précédent papier du *Courrier des Statistiques*<sup>1</sup> qui portait sur l'arbitrage entre coût et qualité en matière de production statistique. On y montrait que la reprise des non-codés, ainsi que la vérification et l'apurement des données, constituent des domaines majeurs de cette production dans lesquels se pose alors la question de l'arbitrage entre coût (ressources humaines utilisées) et qualité (au sens de la fiabilité). Le problème que nous étudions ici est celui de la mise en place effective d'indicateurs de qualité en vue de piloter un tel arbitrage, soit en matière de production statistique, soit dans le domaine de la gestion de référentiels. Sur quelles bases les définir ? À quoi faut-il penser pour les rendre opérationnels ? Comment les insérer dans l'organisation existante, et enfin, quels sont les obstacles à surmonter ?

Une grande difficulté en matière de qualité vient du fait qu'elle oblige à changer de point de vue, de vision des choses. Dès que l'on se réfère à la littérature existante, on peut être surpris des définitions employées. La qualité, selon cette littérature, c'est d'abord l'aptitude à l'emploi, c'est l'aptitude à satisfaire des besoins exprimés ou implicites. Ce n'est donc pas, comme on pourrait le penser, la conscience professionnelle. On peut tout à fait être consciencieux dans son travail sans que, *in fine*, le résultat soit de qualité. Fondamentalement, adopter une démarche qualité implique donc de repenser les pratiques ; cela ne sous-entend pas une critique de la compétence acquise, ni une remise en question de la volonté de bien faire des agents. Il est tout à fait possible d'avoir une organisation dans laquelle chacun est compétent, est animé des meilleures intentions, sans que le résultat global ne soit approprié. Dans une démarche qualité, il s'agit plutôt de sortir d'une logique dans laquelle on s'auto-juge pour passer à une logique d'évaluation objective fondée sur des critères prédéfinis, liés eux-même à un besoin, à un objectif. Ainsi, décréter que les données d'une enquête sont de « bonne qualité » n'a pas grand sens en soi : tout dépend de ce que l'on veut en faire, autrement dit de l'utilisation à venir.



Ce qui caractérise aussi toute démarche qualité, c'est qu'elle doit être vue comme un processus continu, lent, régulier. Obtenir une forme de reconnaissance à un moment donné, via une certification qualité par exemple, ne constitue nullement une fin en soi. Car une photographie de la situation à un instant donné ne prouve en rien que cette qualité se perpétuera. La démarche qualité, les indicateurs de qualité, doivent donc avant tout s'intégrer dans le travail quotidien.

Une telle approche doit naturellement pousser les concepteurs à objectiver le travail de production, à se poser dès le départ la question de l'usage<sup>2</sup> des statistiques (ou des référentiels), mais aussi, on le verra, à repenser leur organisation. Entre autres

choses, il ne faut jamais oublier cette maxime classique des penseurs de la qualité : « la sur-qualité constitue en soi de la non-qualité »<sup>3</sup>. Le fait de trop bien faire peut avoir un impact négatif : coûts additionnels, mais aussi rallongement des délais, qui constituent une composante de la qualité en statistique. Pour éviter ces dérives il est important que les concepteurs objectivent en permanence ce qu'ils souhaitent obtenir, ce qu'ils entendent par « qualité ». Cela requiert la mise en place d'un ensemble d'indicateurs de qualité, coordonné pour l'ensemble de l'organisation, et mis à disposition de tous. Il est donc indispensable que les indicateurs de qualité soient compréhensibles, lisibles et faciles à calculer.

\* Pascal Rivière est inspecteur général de l'Insee, chef du Département des applications et projets de la Direction générale. Il a publié plusieurs articles parus notamment dans le *Courrier des statistiques*, qui sont cités ici.

1. « Approche coût-qualité pour l'amélioration des processus de production statistique », *Courrier des statistiques* n° 105-106, juin 2003, p. 65 à 75.

2. À noter que cette question de l'usage, et son caractère central pour la démarche qualité, sont très présents dans l'article « Qualité et statistique », *Courrier des statistiques* n° 90, juin 1999, p. 47 à 58.

3. En soi, ceci peut paraître paradoxal : l'idée, lorsqu'on parle de qualité, est de rechercher ce qui est juste nécessaire par rapport au besoin, et pas plus. Chercher à aller trop loin, viser une forme d'excellence à tout prix est néfaste, c'est un point sur lequel les spécialistes de la qualité insistent beaucoup.

Dans ce qui suit, nous rappellerons d'abord le contexte, étroitement lié au précédent article. Puis nous proposerons une trame pour la conception d'un système d'indicateurs, avec une démarche allant du général (grandes dimensions de la qualité) au particulier (principes d'indicateurs de qualité). Dans la partie suivante, des pistes seront proposées pour construire un indicateur opérationnel, calculable. Se posera ensuite la question de l'insertion de cet indicateur dans l'organisation, insertion nécessaire pour arriver à une véritable industrialisation de la production et de l'emploi des indicateurs de qualité. Nous verrons pour finir qu'il existe des obstacles non négligeables qui rendent une telle organisation délicate à mettre en place.

### Le contexte : de quelle qualité parle-t-on ici ?

La qualité en statistique recouvre de nombreux aspects, et l'on va se concentrer ici sur une partie d'entre eux. Ainsi ne va-t-on pas parler de « qualité de service » (par exemple la qualité de l'accueil), ni de qualité des publications (cf. les critères de lisibilité, d'accessibilité tels qu'ils sont définis par Eurostat).

On s'intéressera en revanche à la qualité dans la production de données, dans une optique de maîtrise de cette production. On peut distinguer deux cas : la production statistique



4. Par référentiels, on entend par exemple le répertoire d'entreprises et d'établissements SIRENE, la base de répertoires de personnes physiques (BRPP), ou le référentiel géographique.

5. Cela dit, l'utilisation du répertoire d'entreprises à des fins administratives constitue une spécificité française.

6. Note de la rédaction : *tonneau des Danaïdes* : mythologie grecque. Les Danaïdes étaient les cinquante filles du roi d'Argos qui, toutes à l'exception d'Hypermnestre, tuèrent leurs époux durant la nuit de leurs noces. Elles furent condamnées, dans les Enfers, à remplir d'eau un tonneau sans fond. Synonyme : un travail sans fin.

et la gestion de référentiels<sup>4</sup>. On a vu dans l'article précédent que l'essentiel du travail de production statistique résidait dans la vérification et l'apurement des données, ainsi que dans la reprise des échecs du codage automatique. La même remarque peut être faite pour la production de référentiels : l'essentiel de la charge de travail et du délai induit se situe dans la reprise de ce qui n'a pu être traité automatiquement. Des similitudes existent donc, mais ce qui singularise la gestion de référentiels, c'est entre autres choses les deux points suivants : d'une part le caractère continu de la production (il n'y a pas une campagne d'enquête à un moment donné), et d'autre part l'utilisation des données à des fins autres que statistiques<sup>5</sup>.

Nous nous plaçons donc ici dans le cadre d'*indicateurs de production*, appelés à réguler le travail de vérification-apurement et à le rationaliser sur la base du fameux triptyque coût-qualité-délai. Notre propos ne portera donc pas sur les indicateurs de qualité *des statistiques* en général : nous nous concentrerons sur les indicateurs qui permettent de mieux maîtriser la production. Car il s'agit bien d'un véritable *tonneau des Danaïdes*<sup>6</sup> : dans ce fastidieux travail de vérification, on ne sait jamais bien où et en fonction de quoi s'arrêter, dans quel ordre vérifier, et dans quel but. Or l'objectif est bien de se servir d'indicateurs pour prendre des décisions importantes : arrêt des vérifications, attribution de moyens supplémentaires, acceptation d'une qualité moindre sur certaines parties, etc. On souhaite ainsi réguler, mais également savoir où l'on en est, apporter une base solide, rigoureuse, à la notion de « données de bonne qualité », fondée sur l'usage.

On a vu dans l'article précédent que cette notion dépendait des variables et des sous-populations considérées, mais intuitivement on a tendance à se représenter les choses sur la base de la proportion d'erreurs observées : plus elles sont nombreuses, plus on sera amené à penser qu'il existe un problème de fiabilité. Il se trouve que cette vision

est nécessairement biaisée dans l'analyse qu'en font les gestionnaires, sans qu'ils n'y puissent rien : ils ne voient en effet, par construction, que les données « en reprise » (c'est-à-dire pour lesquelles le traitement automatique n'a pas suffi), donc celles qui posent problème, les anomalies, alors qu'il y a peut-être une proportion très importante de données traitées automatiquement et qui ne posent aucun souci. Ils ont alors l'impression que les données contiennent une grande quantité d'erreurs sans que ce soit nécessairement objectif, quantifié. On voit donc toute l'importance d'une mesure systématique, dans laquelle on tente au passage d'objectiver ce que signifie le terme « erreur ». On pressent également l'intérêt d'un ciblage par variables, par sous-population, car il sera utile de déterminer précisément quelles sont les zones les moins fiables de la base de données.

Soulignons que dans notre analyse nous distinguons nettement *indicateurs de qualité* et *indicateurs de suivi*. Lorsqu'on parle de taux d'avancement des vérifications, de temps passé, de nombre de cas restant à traiter, il s'agit d'indicateurs de suivi de la production : ils répondent à une logique totalement interne et ne constituent à aucun moment des indicateurs de qualité, car ils ne permettent aucunement de visualiser l'impact de l'avancement de la production sur la fiabilité des résultats statistiques en sortie. Autrement dit, on ne s'y pose pas la question de l'usage final des données.

### Concevoir un système d'indicateurs pour la production : une démarche descendante

La mise en place d'indicateurs de qualité constitue en pratique un travail très difficile. Vingt fois sur le métier il faut remettre l'ouvrage. La difficulté n'est pas tant d'imaginer quelques indicateurs ; cela déjà n'est pas simple mais il est toujours possible de construire de telles données.

Le problème est de s'assurer que l'ensemble des indicateurs que l'on construit soit cohérent et relativement exhaustif, mais aussi qu'il assure une compréhension partagée des objectifs, de ce que les concepteurs veulent faire... et que les indicateurs qui composent cet ensemble soient calculables en pratique. Il ne s'agit pas de faire assaut d'habileté méthodologique, c'est plutôt l'inverse : faire simple et efficace. Ne négligeons donc pas le danger de la liste d'indicateurs et des multiples tableaux de chiffres. Enfin, n'oublions pas que lorsqu'on a réalisé cela, on n'a fait qu'une petite partie du travail, car tout est dans la synthèse (présentation, commentaires, etc.)

Comment démarrer ? Surtout, éviter de s'engager tout de suite dans la définition des indicateurs eux-mêmes : c'est une impasse, qui conduit à perdre d'entrée le recul nécessaire. Il est important de raisonner en respectant les différents niveaux d'abstraction. On devra partir « du haut », *i.e.* des principaux concepts, en descendant petit à petit dans l'opérationnel et en respectant à chaque étape l'étanchéité entre les niveaux d'abstraction.

Concrètement, si l'on prend le cas d'un répertoire d'entreprises par exemple, on va commencer par décrire les *principaux sujets qui caractérisent la fiabilité du répertoire* : par exemple, sur-couverture (veiller à ce que les unités présentes soient à leur place), sous-couverture (veiller à ce qu'on n'en oublie pas), fiabilité des variables elles-mêmes. À ce niveau-là, il est crucial d'être exhaustif.

À partir de là, on pourra *distinguer des catégories*. En matière de sur-couverture, par exemple, on pourra isoler une catégorie « doubles comptes », correspondant aux unités qui apparaissent deux fois dans le répertoire. On pourra également avoir une catégorie regroupant les faux actifs. Là aussi, on veillera à ne pas en oublier, en définissant au besoin des sous-catégories : c'est en quelque sorte une nomenclature que l'on cherche à se constituer.

*In fine*, une fois les différentes catégories et sous-catégories déterminées, l'idée sera de bâtir non pas des indicateurs mais des *principes d'indicateurs* : il s'agit par exemple, pour un répertoire d'entreprises, du taux de NPAI (unités non présentes à l'adresse indiquée) ; pour une enquête, de l'erreur d'échantillonnage, du taux de non-réponse, etc. Dans ces exemples, on voit bien que l'on reste vraiment au niveau du principe, car on ne donne, à ce stade, aucun élément sur la manière dont l'indicateur opérationnel va être défini. Lors de cette étape, il ne s'agit donc aucunement de s'interroger sur l'existence même de données permettant de calculer l'indicateur, puisqu'on ne connaît pas le détail des structures. On reste encore dans l'abstrait, en définissant simplement ce qu'il serait souhaitable d'avoir... quitte à constater, ultérieurement, que l'on n'a pas la possibilité de fabriquer un véritable indicateur correspondant au principe souhaité.

L'étape suivante est d'une toute autre nature : ayant des principes d'indicateurs, on va essayer de passer à des indicateurs opérationnels. Cela signifie que l'indicateur sera défini sous la forme d'une formule mathématique applicable à partir d'un certain nombre de variables disponibles (existantes ou que l'on fabrique).

### **Vers des indicateurs opérationnels : champ, variable d'intérêt, domaines-cibles...**

Poursuivons notre démarche descendante. Étant « partis du haut », nous sommes arrivés à des principes d'indicateurs, encore vagues. À principe d'indicateur donné, reste maintenant à décrire comment passer à un indicateur concret. L'idée de base est de distinguer systématiquement quatre dimensions dans un indicateur<sup>7</sup> : son champ, la variable d'intérêt sur laquelle l'attention se focalise, l'ensemble des domaines cibles pour lesquels il sera calculé, et enfin son mode de calcul.

Le *champ* auquel s'applique l'indicateur (temporel ou spatial) n'est rien d'autre que la définition de la population à laquelle se réfère l'indicateur, à la fois en termes de date (ou de période de référence) et d'unités statistiques (ménages, individus, entreprises, établissements) : par exemple les entreprises du secteur industriel au 31 décembre 2004. Il va de soi qu'un indicateur se réfère à une *variable d'intérêt* donnée, existante (effectif d'une entreprise, statut actif/non-actif) ou déduite (le fait qu'une unité soit en double compte, par exemple).

La grille des *domaines-cibles* (ou sous-populations cibles) constitue quant à elle un élément vital pour piloter la démarche qualité, qui impose d'avoir préalablement défini les objectifs associés à la production : plus ces domaines seront fins, nombreux, plus les coûts seront élevés, plus il sera nécessaire d'effectuer des arbitrages entre domaines. Bien entendu, l'idée générale est de s'appuyer systématiquement sur les besoins des « utilisateurs ». Mais il apparaît en pratique que les utilisateurs en question sont extrêmement variés, pas toujours faciles à cerner, et très gourmands en matière de données. Il s'avère difficile de se fonder à 100 % sur un ensemble disparate de besoins : sur la question du niveau de finesse des statistiques, le statisticien aura ainsi pour charge de trouver un bon compromis entre la masse des besoins (ou du moins ce qu'ils révèlent du niveau de finesse attendu) et ce qui est réalisable avec les moyens disponibles, le tout afin de définir l'ensemble de domaines-cibles.

### **... et surtout un mode de calcul applicable**

Une fois le champ, la variable et les sous-populations cibles fixés, reste la

7. Pour plus de précisions, cf. « Cadre général pour la mise en place d'indicateurs de qualité dans un répertoire d'entreprises », note 127/E210, Insee, août 2000. On peut en trouver une version en anglais dans les actes de la conférence « Roundtable on business survey frames », Auckland, novembre 2000, disponible sur le site de Statistics New Zealand.

principale dimension de l'indicateur : son *mode de calcul*. À ce sujet, notons que le terme « indicateur » n'est pas choisi au hasard car une véritable mesure de la qualité ne serait pas atteignable. C'est une remarque désormais classique : s'il était possible de mesurer parfaitement la fiabilité d'un fichier X en le comparant à un fichier de référence Y... alors le fichier Y se substituerait au fichier X, dont la production perdrait tout intérêt. Dans le domaine des données statistiques et administratives, il n'y a pas en général de fichier de référence absolu, indiscutable, sans aucune erreur<sup>8</sup>. Les indicateurs que nous bâtissons doivent systématiquement tenir compte de cela. Il nous faut donc trouver des chemins de traverse qui permettront simplement d'approcher le principe d'indicateur souhaité.

Prenons un exemple : comment représenter l'idée de fiabilité de l'activité principale d'une entreprise ? Dans la mesure où il n'existe pas *a priori* de référence externe, une possibilité simple est de prendre en compte l'ancienneté de cette variable, c'est-à-dire le nombre d'années depuis lequel elle n'a pas été modifiée. Il est clair que plus l'activité principale est ancienne et moindre sera la qualité de cette variable. C'est une information partielle certes, mais importante, qui se fonde sur des données disponibles (date de validité). Une autre possibilité, lorsqu'il n'y a pas de base de référence, consiste à en créer une, à constituer l'information : aller vérifier la donnée (appel téléphonique, recherches sur Internet ou Minitel, enquête sur le terrain), ici l'activité économique, sur un échantillon. Cela pose deux problèmes : le coût (pour

cette raison, on limitera peut-être le champ d'investigation), et la fiabilité de cette nouvelle mesure (il faudrait idéalement un expert, capable de déterminer à chaque fois la « bonne » activité principale).

Derrière ces remarques, on voit se profiler une simple idée de bon sens : pour construire un indicateur de qualité opérationnel, on ne peut que se fonder sur des données disponibles, ou sur des informations additionnelles que l'on est capable de se procurer. La faiblesse de ces informations conduit parfois à des solutions insatisfaisantes intellectuellement... ou à des solutions chères si l'on veut absolument mener une enquête spécifique.

Pour être plus précis, le mode de calcul de l'indicateur se passe souvent en deux temps. On détermine d'abord une sorte d'indicateur de qualité individuelle, donnée par donnée (ou unité statistique par unité statistique) : écart à une donnée supposée de référence, âge de la donnée, mesure de cohérence en fonction d'une quelconque norme<sup>9</sup>. Il peut aussi s'agir de données binaires : le fait que la donnée ait ou non été redressée, manquante, en anomalie, vérifiée par un questionnaire, modifiée par un questionnaire, jugée conforme à l'avis d'un expert, le fait que l'unité soit présente à l'adresse indiquée, qu'elle ait répondu, etc.

Dans un deuxième temps, on effectue une synthèse de ces informations, *le calcul se faisant dans le champ que l'on s'est fixé au départ, et sur chaque domaine-cible*. Lorsqu'on part d'un élément d'information individuel binaire, par exemple le fait d'être en non-réponse ou pas, l'indicateur de qualité sera une proportion : dans ce cas, le taux de non-réponse, qui ne fait qu'agrèger ces indicateurs binaires. Cette proportion pourra être simple (taux de réponse à une question, taux d'unités redressées, taux de NPAI, taux de conformité, taux de doubles comptes...) ou faisant appel à plusieurs variables binaires

(taux d'unités modifiées parmi celles qui étaient en anomalie).

Si les indicateurs individuels sont quantitatifs, on se référera à leur distribution, avec un indicateur de tendance centrale (moyenne, médiane, maximum, quantiles). Par exemple : délai maximum de prise en compte des naissances dans le répertoire, âge moyen d'une activité principale, neuvième décile des « notes de cohérence ». On peut également s'appuyer sur des indicateurs de dispersion (coefficients de variation, le plus souvent), qui utiliseront non pas une distribution d'indicateurs de qualité individuels, mais bien les données elles-mêmes, ainsi que leurs poids de sondage et d'autres caractéristiques (caractère répétant ou non).

### Cas des référentiels, cas de la production statistique : quelques pistes

Dans le cas de la gestion de référentiels, on peut distinguer cinq façons de procéder pour construire les indicateurs individuels :

- utiliser des sources d'information externes auxquelles on se comparera (sources administratives, enquêtes...) : l'indicateur individuel est alors, pour chaque donnée, une mesure d'écart à une source supposée être de référence (avec toutes les réserves que cela suscite) ;
- effectuer une investigation spécifique, de type enquête d'amélioration du répertoire, et là aussi se fonder sur l'écart vis-à-vis de cette nouvelle source d'information, donnée par donnée ;
- analyser la cohérence interne par rapport à d'autres variables du répertoire (exemple : prendre un ratio chiffre d'affaires/effectif) : l'indicateur de qualité est alors qualitatif, selon l'importance de l'incohérence ;
- analyser la cohérence temporelle se référant à la valeur de la variable lors d'une période précédente (année, mois, trimestre) : on peut

8. Cf. à ce sujet I. Boydens, « Informatique, normes et temps », Bruylant 1999.

9. Ceci peut amener à un indicateur qualitatif prenant en compte la cohérence supposée avec plusieurs autres variables. Par exemple, pour avoir un indicateur de fiabilité individuel de la variable chiffre d'affaires (CA) lors de la période n, on va calculer pour chaque unité le ratio  $CA(n)/CA(n-1)$ , ainsi que le ratio  $CA(n)/\text{Effectif}(n)$ , Valeur ajoutée (n)/CA (n). Parallèlement à cela, on déterminera des intervalles de validité pour ces ratios, en analysant leur distribution, domaine par domaine. On attribuera à chaque entreprise, pour la variable chiffre d'affaire, une « note » qui sera fonction du nombre de cas où le ratio appartient à l'intervalle de validité.

d'ailleurs combiner cohérence interne, temporelle et inter-sources pour synthétiser le tout au sein d'une seule et même « note » ;

– s'appuyer sur une métadonnée individuelle : on pense notamment à la *date de validité de l'information*, qui permet de déterminer la fraîcheur d'une donnée, ou toute *date de mise à jour*, qui permet de mesurer par exemple un délai de prise en compte d'une naissance dans le répertoire.

Dans le cas de la production statistique<sup>10</sup>, il existe des indicateurs de qualité fondés sur la cohérence interne au questionnaire, sur la cohérence temporelle, ou sur l'écart à d'autres sources (cas du codage automatique où l'on peut faire appel à un expert). De fait, ce sont souvent ces indicateurs de cohérence individuels qui vont permettre, en production, de déterminer sur quelles unités le gestionnaire doit intervenir. Les indicateurs fondés sur les métadonnées (ici les données de gestion) ne sont pas rares non plus : le taux de réponse (à une question, à l'ensemble du questionnaire) est une donnée utile, qui pourra par la suite conduire à modifier les libellés de questions.

On peut utiliser également des indicateurs de précision, prenant en compte différents phénomènes (erreur d'échantillonnage, erreur de mesure, erreur due à la non-réponse...)<sup>11</sup>. L'expérience montre que de tels indicateurs (coefficients de variation notamment) sont très rarement utilisés pour piloter la production<sup>12</sup>. Le point important en la matière, est de ne pas chercher obligatoirement à bâtir un indicateur de précision sophistiqué, satisfaisant méthodologiquement : rappelons qu'en effet, on ne vise pas les indicateurs de qualité « finaux », mais bien des instruments de pilotage de la production. Ainsi, on peut tout à fait calculer la précision très approximativement en raisonnant uniquement sur la variance d'échantillonnage, en « faisant comme si » l'échantillon se limitait au sous-ensemble des répondants, ce qui exige de réajuster les poids de sondage en fonction.

### Mise en pratique dans l'organisation : pérennité, lien qualité → action...

Supposons maintenant que nos indicateurs aient été définis et qu'ils soient calculables opérationnellement. Comme la démarche qualité n'a d'intérêt que si elle peut faire l'objet de comparaisons dans le temps et de pérennité, nos indicateurs ne seront pas déterminés une seule fois, ponctuellement, mais bien de façon systématisée. Les calculer régulièrement, industrialiser leur production requiert la mise en place de toute une infrastructure : une documentation méthodologique, fût-elle légère, une application informatique stable et dont la maintenance est assurée, et surtout une organisation du travail clairement définie (répartition des rôles et responsabilités, échanges d'information entre équipes, accompagnement par des formations). Cette infrastructure aura également vocation à traiter de la même façon les indicateurs de suivi de production (nombre de questionnaires traités, temps passé...), qui reflètent un état d'avancement dans le fonctionnement interne.

Ce n'est pas tout : il n'est pas imaginable de construire des indicateurs de qualité s'il est impossible de les améliorer. Dans ce dernier cas, en effet, ils ne présenteraient pas d'intérêt et ne serviraient qu'à mettre en évidence un défaut structurel, sans possibilité d'intervention : l'indicateur de qualité deviendrait ainsi un indicateur de culpabilité ! Pour chaque indicateur, il est donc important de veiller à ce qu'il y ait toujours une action de correction susceptible de l'améliorer. Tout indicateur dont la valeur est jugée mauvaise devrait en principe pouvoir déclencher une action, à court ou moyen terme. Ceci vaut pendant la phase de production, dans le cas d'une enquête : ainsi, constater, *via* des indicateurs de précision, que les statistiques de tel secteur d'activité sont encore très imprécises, devrait conduire à mobiliser plus de gestionnaires sur

les tâches de vérification-apurement des questionnaires de ce secteur. L'action peut également avoir lieu entre deux campagnes : ainsi, dans le recensement, des indicateurs de qualité du codage de la profession pourront permettre de repérer les catégories socioprofessionnelles les moins bien traitées par le chiffrement automatique ; on mettra alors l'accent sur celles-ci dans la mise à jour des bases de connaissances « profession » utilisées pour le codage automatique, en vue de mieux chiffrer cette variable lors de la campagne suivante. Dans le cas d'un répertoire, où la mise à jour est permanente et non pas effectuée par campagnes, les actions sont de nature différente : elles consisteront en des EAR (enquêtes d'amélioration du répertoire), que l'on pourra bien cibler sur les variables d'intérêt et les sous-populations les plus critiques.

### ... mais aussi lisibilité

Afin que l'enchaînement *indicateur « mauvais » → action corrective* soit efficace, mais aussi dans un but d'appropriation collective et de pérennité, la *lisibilité* de ces données synthétiques sur la qualité s'avère essentielle. Elle requiert au préalable un lent processus d'explicitation, une pédagogie adaptée : des réunions pour confronter les points de vue, des exposés, tout un accompagnement à travers des formations notamment, et bien sûr une documentation adéquate. Au sein des rapports qualité eux-mêmes, cette lisibilité passe par une représentation standardisée, des titres compréhensibles, une comparaison avec le jeu d'indicateurs précédents, une grande vigilance sur les titres, la vérification des valeurs qui paraissent

10. Il existe toute une batterie d'indicateurs de qualité, définis à la fin des années 90 par Eurostat dans le cadre du groupe de travail sur la qualité des statistiques, et dont certains peuvent être utilisés pour le pilotage de la production.

11. Cf. encadré de l'article de Philippe Brion « La prise en compte de la qualité dans les enquêtes auprès des entreprises », dans ce même dossier.

12. Pour un essai en la matière, cf. « Impact on mean squared error as a score to handle data editing », Insee, document de travail DSE n° 2002/10.

aberrantes, l'élagage de ce qui n'apporte pas d'information, un souci de simplification permanente, mais aussi des commentaires d'accompagnement des séries d'indicateurs.

Le danger serait en effet de ne produire que des tableaux de chiffres, sans travail sur la forme, sans texte explicatif, en considérant implicitement que « ce n'est pas la peine, tout le monde comprend ». Il s'agit là d'un problème classique, fréquent, révélateur d'une logique interne, à l'opposé d'une démarche qualité. On n'insistera jamais assez sur ce point : lorsque tous les indicateurs de qualité sont calculés, produits, ce n'est pas fini. Car l'essentiel du travail reste à venir : agencer l'ensemble des indicateurs de façon synthétique et compréhensible dans un seul et même document. Il s'agit là d'un investissement permanent, et non ponctuel. Par conséquent, si les rapports ou bilans qualité ont vocation à une certaine stabilité dans le temps (dans une optique de comparabilité), ils ne sont pas non plus gravés dans le marbre : on cherchera donc à les faire évoluer, par enrichissement ou simplification, afin qu'ils soient aussi lisibles et efficaces que possible<sup>13</sup>.

### Quels obstacles ?

La mise en place d'indicateurs de qualité nécessite une formalisation poussée des objectifs. Il se trouve simplement que ces derniers sont souvent peu explicites et qu'ils peuvent se révéler excessivement variés, donc nombreux. Définir *a priori* et afficher un jeu d'objectifs réduit constitue déjà une difficulté

13. Par exemple, la structure du bilan qualité des enquêtes auprès des entreprises a évolué, notamment dans sa partie « précision » : dans la version initiale, on s'intéressait à la distribution des coefficients de variation dans l'ensemble des domaines-cibles, pour chaque variable, et l'on demandait la médiane, le neuvième décile et le maximum de cette distribution. C'était finalement assez illisible ; on a remplacé cela par l'ensemble des coefficients de variation pour tous les domaines-cibles.

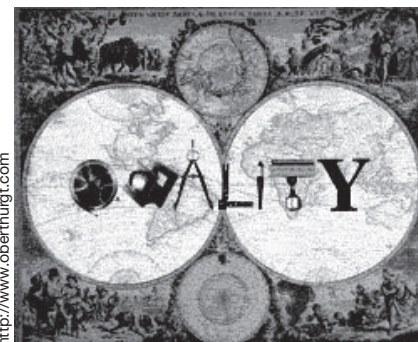
énorme : au fond, chacun pense que les données peuvent servir à tous les usages, et le fait de poser une limite dans ces usages requiert un engagement fort, qui fait inévitablement des mécontents. À objectifs donnés, on ne peut pas non plus avoir des centaines d'indicateurs : le « bon » rapport qualité est synthétique, et le fait d'omettre de nombreux indicateurs intéressants constitue en pratique un véritable crève-cœur pour les concepteurs, car ils devront élaguer ce qu'ils auront eux-mêmes conçu.

Dès lors que les indicateurs existent, il faut aussi savoir les faire accepter, en tant que partie intégrante du travail. De ce point de vue, l'expression « opération qualité » peut être révélatrice de dysfonctionnements car elle est inadaptée à double titre : d'une part, elle suggère qu'il s'agit de quelque chose de ponctuel, or une démarche qualité n'a de sens que dans la durée, dans l'amélioration continue et régulière ; d'autre part, cette expression laisse supposer que ce travail sur la qualité est séparé du processus de production, qu'elle n'en fait pas naturellement partie, or le propre d'une démarche qualité est de servir de fédérateur, et surtout pas de tâche supplémentaire, de « cerise sur le gâteau ». Ainsi, la clé de la réussite des indicateurs de qualité réside-t-elle dans leur intégration au fonctionnement courant, dans une véritable dilution au sein de la production : tant que l'on parlera d'opération qualité (en tant qu'opération « coup de poing »), tant que les bilans qualité d'enquête seront vécus comme un supplément de travail et non une composante indispensable de celui-ci, on ne progressera pas. L'idée de *démarche*, de régularité et de pérennité, est bien essentielle et caractéristique de l'approche qualité. Elle va à l'encontre du côté spectaculaire et vain d'une *opération*, qui donnerait l'illusion que l'on peut « nettoyer » les données et améliorer sensiblement les choses via une intervention ponctuelle.

Cela signifie concrètement que les objectifs fixés à la production (répertoires, enquêtes, traitement de sources externes), ainsi que les indicateurs de qualité associés, doivent être réellement portés par tous les acteurs concernés, de haut en bas de l'organigramme, avec une implication forte de la hiérarchie. Notons que la LOLF (Loi organique relative aux lois de finances), qui impose progressivement en France une approche par indicateurs de performance, peut s'avérer de ce point de vue un véritable catalyseur pour la démarche qualité.

Dans le même temps, n'oublions pas que les indicateurs et leur suivi ne doivent pas faire négliger l'organisation, les savoirs faire et la culture existants, qui privilégient, à l'Insee comme dans d'autres instituts nationaux de statistique, une grande minutie. Il ne s'agit évidemment pas de transformer radicalement les façons de travailler, mais plutôt d'y instiller progressivement un autre angle de vue, une prise en compte plus marquée et systématique des objectifs de qualité, ainsi que de la quantification de ceux-ci : en d'autres termes, une évolution de la « conscience professionnelle » de chacun, plus ouverte sur l'extérieur et l'utilité de ce que l'on produit.

Plus elle sera incorporée dans les mœurs, digérée, plus la démarche qualité progressera : son succès résidera paradoxalement dans sa banalisation, dans la quasi-disparition du terme.



<http://www.oberthurgt.com>