2.2 Concrétisations

2.2.1 0 D Q G D W G ¶ L Q I R U P D W L R Q

Pour donner suite à ce qui a été décrit ci-dessus, la statistique publique est tenue de recourir à de nouvelles sources de données permettant de remplir son mandat fondamental de couvrir les besoins en informations V W D W L V W L T X H V S H U W L Q H Q W V G ¶ L Q W ¶ U ê W S X E O L F V G D Q V O H V S U L Q F L S H V fondamentaux et de la Charte. Dans le CoP, la pertinence des informations statistiques figure dans le S U L Q F L S H O H S U L Q F L S H p W D Q W F R Q V D F U p j O ¶ L Q G p S H Q G D Q F H G

2.2.2 Les codes de conduite ne contiennent pas de directives sur les sources de données

Des contraintes telles que par exemple des restrictions relatives aux sources de données sont-elles formulées dans les codes de conduite considérés?

- Dans les Principes fondamentaux, le principe 5 indique que les données utilisées à des fins statistiques peuvent être tirées de toutes sortes de sources.

- Le préambule de la version révisée en 2017 du Code de bonnes pratiques mentionne entre autres que © / ¶ R E M H F W L I G H O D Z D J U H O H V F K D Q J H P H Q W V W O H V L Q Q R Y D W L R Q V O production et la diffusion des statistiques officielles dans le système statistique européen et au-delà, par H [H P S O H O H V Q R X Y H O O H V V R X U F H V G ¶ L Q I R U P D W L R Q V O ¶ P R G H U Q L V D W L R Q G X F D G U H M X U L G L T X H W O H V U p V X O W D W V G H V code de bonnes pratiques.». La référence à des «informations provenant de multiples sources de données» dans la version révisée est reprise dans le principe 2 ainsi que dans plusieurs indicateurs du Code de bonnes pratiques.

- La Charte ne prévoit aucune restriction relative aux sources de données. À la différence des Principes I R Q G D P H Q W D X [H W G X & R G H G H E R Q Q H V S U D W L T X H V H O O H Q ¶ p Y R des sources de données multiples. La Charte devrait ê tre complétée en conséquence lors de la prochaine révision.

2.3 Conclusion

Du point de vue des trois codes de conduite considérés, le recours à des sources de données diverses ne pose aucun problème fondamental. Mais le traitement de nouvelles grandes sources de données soulève des questions à différents niveaux, que nous évoquons brièvement ci-après (sans garantir leur exhaustivité):

3. Nouvelles questions

3.1 Acquisition de données auprès de fournisseurs privés

Comme les fournisseurs privés ne sont le plus souvent pas tenus légalement de remettre des données à la statistique publique, ils le font sur une base volontaire.

3.1.1 * D U D Q W L H G H O ¶ L Q G p S H Q G D Q F H S U R I H V V L R Q Q H O O H

Pour la statistique publique, il est primo U G L D O G ¶ D V V X U H U O ¶ L Q G p S H Q G D Q F H S U R I H V V faisant preuve de transparence dans la désignation et le choix des fournisseurs de données. En F R Q W U H S D U W L H O H V V H U Y L F H V V W D W L V W L T X H V G e s r m t e u r s d e d o n n é e s V X U H U exclusivement à des fins statistiques.

3.1.2 Question des coûts

, O F R Q Y L H Q W E L H Q V € U G H G L V F X W H U G H O D T X H V W L R Q G H V F R € W V données de fournisseurs privés. Cet aspect financier touche des problématiques politiques et éthiques, notamment des questions liées à O ¶ p W H Q G X G H U Y L F H S X E O L F G D Q V O H G R P D L Q H G H données et de son utilisation entre autres par la statistique publique.

3.2 Un point particulier à considérer: information des milieux interrogés en cas de recours à de nouvelles sources de données?

Le recours à de nouvelles sources de données telles que les réseaux sociaux ou les données issues de la W p O p F R P P X Q L F D W L R Q Q ¶ H V W p v u r l a S t a t i s t i q u e p u b l i q u e , d u m o i n s p a s d i r e c t e m e n t . L e s milieux interrogés, soit ici les exploitants / propriétaires des sources de données concernées, doivent être informés des E X W V H W G H V E D V H V O p J D O H V H W F G H V H Q T X r W H V H Q Y H U W X G H (base légale). 8 Q H W H O O H L Q I R U P D W L R Q Q ¶ H V t - à - d i r e p a s d e s t a n d a r d s t a t i s t i q u e U E O p P D V X I I L V D Q W G ¶ L Q f o r m e r l e s d é t e n t e u r s d e d o n n é e s que leurs données personnelles sont communiquées à la statistique publique à des fins statistiques via les Conditions générales (CG)? Doivent-ils absolument en être informés? Ou faut-il limiter la communication à des données non personnelles?

3.3 Nouveaux défis méthodologiques et qualitatifs

Les nouvelles grandes sources de données ne sont généralement pas conçues à des fins statistiques. Elles ne correspondent pas à la statistique publique ni en termes de volume ni à ses modèles traditionnels au niveau de la structure ou du contenu. Les nouvelles sources de données ne répondent pas nécessairement aux exigences statistiques en ce qui concerne les définitions, les standards, la représentativité, la couverture ou la continuité. Cette situation amène donc de nouvelles questions et possibilités quant aux méthodes statistiques et aux exigences qualitatives; elles sont abordées actuellement dans le cadre de la stratégie G ¶ L Q Q R Y D W L R Q V X U O H V G R Q Q p H V G H O ¶ 2) 6 p x e m p l e p o u r l ' a n a l y s e d e q u a l i t é d e V R Q W données dont le service statistique ne sait pas comment elles ont été créées et pour lesquelles les P p W D G R Q Q p H V V ¶ l n t e r p r è t e n t p a r f o i s d i f f i c i l e m e n t . Il faut donc renforcer les divisions de recherche dans la statistique publique ainsi que la collaboration avec les hautes écoles et instituts de recherche; la division DSSM nouvellement créée tient compte de ces enjeux.

3.4 Labellisation des résultats «en dehors» de la production statistique

'DQV OH FDGUH GH OD VWUDWpJLH G¶LQQRYDWLRQ VXU OHV GRQQpHV adéquation avec les directives de la Charte et peuvent dès lors être intégrés dans la production statistique. /RUVTXH FH Q¶HVW SDV OH FDV VL S H[OD TXDOLWp QH FRUUHVSR se trouve seulement à la phase de développement de projets, il faut chercher des possibilités répondant aux besoins des parties prenantes sans compromettre la crédibilité de la statistique publique. La voie JpQpUDOHPHQW VXLYLH DFWXHOOHPHQW j O¶2)6 DLQVL TXH GDQV OH publier de tels résultats issus de la statistique expérimentale sous un label et avec les m p W D L Q I R U P D W L R Q rapportant.

3.5 Limites

La production statistique établie en recourant à de nouvelles méthodes doit respecter les directives des codes de conduite valant pour la statistique publique. Quelles solutions devons-nous adopter pour des cas FRPPH «

3.5.1 « OD VWDWLVWLTXH H[SpULPHQWDOH

Dans quelle mesure peut-on outrepasser ces directives, par ex. dans la statistique expérimentale, et FRPPHQW SUpYRLU O¶LQIRUPDWLRQ j FH VXMHW HW OD ODEHOOLVDW

- Est- LO SRVVLEOH G¶DSSRUWHU GHV GpURJDWLRQV DX[GLUHFWLYHV pWDEOLU G¶DXW directives éthiques?

3.5.2 « OHV GRQQpHV DGPLQLVWUDWLYHV HW VWDWLVWLTXHV

Les délimitations entre «statistique ª HW ©QRQ VWDWLVWLTXHª GH PrPH TX¶HQWUH S et «données statistiques» peuvent être floues.

- Peut- RQ IL[HU GHV OLPLWHV HW HQ IQ GH FRPSWH GHV GpILQLWLRQ données?
- Comment procéder en sachant que les informations conçues au départ pour la statistique sont aussi de SOXV HQ SOXV VRXYHQW VROOLFLWpHV SDU O¶DGPLQLVWUDWLRQ et les codes de conduite.

3.6 Autres *(apparaîtront au fil de l'avancement du processus)*

$X QRP GX &RQVHLO G¶pWKLTXH

Peter Laube