

PREFACE

CHARLOTTE VANNESTE ET FRÉDÉRIC VESENTINI

Dans un récent rapport de l'OCDE, William Lehr (Massachusetts Institute of Technology)¹ met en évidence combien les institutions en charge des statistiques sont actuellement confrontées à de nouveaux défis en terme de production et de diffusion de l'information publique, mais également combien ces défis peuvent être transformés en autant d'opportunités. Selon lui, c'est le modèle de fonctionnement traditionnel de ces institutions qui est bouleversé, avec une mission de communication ramenée au centre des enjeux. Lehr décrit également une seconde transformation, liée cette fois à la temporalité du processus statistique traditionnel, en décalage avec les nouvelles attentes, que ce soit en termes de prise de décision ou d'évaluation des politiques publiques. L'auteur de souligner finalement combien l'environnement offert par Internet était à même de permettre de nouvelles pratiques en matière d'utilisation des statistiques publiques, que ce soit par une accessibilité accrue, une alimentation plus souple ou la garantie d'une amélioration continue de la qualité des données.

Les récentes initiatives d'ouverture des données publiques *Open Data* qui fleurissent un peu partout dans le monde, des Etats-Unis à Hong Kong et de l'Irlande à l'Arabie Saoudite, vont très clairement dans ce sens. Elles encouragent les gouvernements à mettre à disposition du plus grand nombre un maximum de données concernant leur territoire. On y trouve pêle-mêle des statistiques démographiques, économiques, sociales, institutionnelles, météorologiques mais aussi, signe des temps, des informations sur la qualité de l'air ou l'état du trafic routier. Pour l'heure, l'initiative est néanmoins plus ou moins investie selon les pays. La Belgique n'est, pour sûr, pas dans le peloton de tête eu égard au nombre de données à disposition. Mais plus que la quantité d'information, c'est peut-être leur utilisabilité qu'il s'agit d'évaluer. Des efforts ont à cet égard été réalisés dans certains pays qui ont mis en place des outils dédiés à l'exploration et la mise en forme des données. Relevons le cas de la France et du logiciel *Data Publica*² qui permet, par exemple, de cartographier les données ou de la Suède et du logiciel *PC-Axis*³, de plus en plus utilisé par les instituts statistiques en Europe.

Sans remettre en question les fondements de ces initiatives, ces efforts de mise à disposition ne suffisent toutefois pas à occulter l'absence complète d'infor-

¹ Lehr, W. (2012), "Measuring the Internet: The Data Challenge", *OECD Digital Economy Papers*, No. 194, OECD Publishing.

² <http://www.data-publica.com/> (consulté le 19 juillet 2012).

³ <http://www.scb.se/pc-axis> (consulté le 19 juillet 2012).

mations critiques sur la production des données. Or, il s'agit là d'un élément essentiel de l'évaluation de la qualité de l'information, sans compter que, paradoxalement peut-être, de nombreuses expériences d'informatisation de l'administration et de la production statistique ont complexifié encore davantage les processus de production. Ainsi dans le champ des données belges relatives à la justice pénale, l'informatisation du processus de production statistique n'a pas permis pour autant de couvrir l'ensemble du champ des besoins statistiques, ni de rencontrer les impératifs de fiabilité, d'exhaustivité et de validité d'une statistique qui de surcroît se devrait « intégrée ». Au contraire, les difficultés de l'informatisation ont entraîné dans leur sillage nombre de problèmes nouveaux sur le plan de la production statistique. L'utilisation de logiciels différents et non compatibles dans les divers secteurs concernés, la prédominance de finalités administratives auxquelles ne sont accolées que secondairement les finalités statistiques, ou encore le caractère volatile de données non stabilisées produites par des systèmes en production sont autant de problématiques nouvelles affectant la production de données statistiques.

En outre, concernant les statistiques pénales, l'ouverture des données prend une dimension toute particulière au vu de l'importance que recouvre, dans le débat démocratique, la question de la criminalité et de son traitement social. Mais pour que cette ouverture soit profitable, il reste un énorme travail à accomplir, tant sur la production des données que sur leur modalité de diffusion. La production de statistiques criminologiques adéquates constitue un premier enjeu fondamental. Une bonne statistique doit en effet permettre à ceux qui sont responsables du fonctionnement d'institutions sociales d'accomplir convenablement leur tâche et d'en rendre compte. Une fonction particulièrement importante dans une démocratie est en effet celle « d'accountability » qui renvoie au devoir incombant aux instances publiques de rendre des comptes et de justifier l'exercice des tâches qui leur sont confiées. La statistique doit y jouer un rôle essentiel. Elle doit également apporter un éclairage sur le fonctionnement de l'appareil de justice et en permettre ainsi l'évaluation : « grâce à cette transparence, le fonctionnement effectif du droit doit pouvoir être mis à l'épreuve de ses propres intentions de principe ». Pourtant même si des avancées essentielles ont été réalisées en Belgique ces dernières années, des lacunes importantes doivent encore être comblées sur le plan de la production, en particulier au niveau de l'intégration des différentes données existantes. Le projet actuel d'intégration des différentes bases de données existantes vise tout particulièrement la dimension d'intégration verticale, entre les différentes phases du fonctionnement du système pénale, qui est à présent encore totalement défaillante. La diffusion de ces données constitue le second enjeu : d'une part, en livrant une information critique et pondérée au plus grand nombre et d'autre part, en rendant possible un débat contradictoire qui suppose l'exis-

tence d'éléments de référence communs aux divers acteurs. L'enjeu de la diffusion est par conséquent d'organiser la transition d'une statistique de gouvernement à une statistique publique. Il s'agit en d'autres termes de rendre aisément accessible d'immenses corpus statistiques, tout en offrant une expertise scientifique et critique quant à la portée et la signification des chiffres.

L'histoire de cet ouvrage rencontre l'histoire de deux projets de recherche qui tout en ayant des objets bien spécifiques et distincts connaissent des intersections multiples et des enjeux communs évidents. Le premier est le projet « Instap », succédant au projet « Quetelet.net », initié par un partenariat universitaire et financé par la Politique scientifique. Son objet est la création d'un outil de numérisation et de diffusion des statistiques pénales belges existantes. Le second est un projet de recherche développé par l'Institut National de Criminalistique et de Criminologie suite à la demande formulée par le Ministre Jo Vandeurzen, en fin 2008, de travailler au développement d'un *datawarehouse* global pour la justice pénale qui permettrait d'articuler entre elles les informations enregistrées dans les différentes banques de données. C'est le démarrage de ces deux projets, axé l'un sur la diffusion des statistiques, l'autre sur leur production, qui a créé l'occasion d'organiser en mai 2009 un colloque international s'adressant en particulier à tous les acteurs belges concernés⁴. L'ouvrage est ainsi constitué des contributions des experts étrangers présents à ce colloque ainsi que des contributions relatives aux deux projets de recherche et qui à l'heure de clôturer cette édition ont tous deux abouti à des résultats.

L'ouvrage présente ainsi sept contributions apportant chacune à leur manière un éclairage sur les statistiques pénales, leurs enjeux de production et de diffusion. La première est celle de **Charlotte Vanneste**, maître de recherches à l'Institut national de criminalistique et de criminologie (INCC), un établissement scientifique de l'Etat dont l'expertise est mobilisée tant au niveau de la production que de l'exploitation des données. Elle y décrit comment on tente en Belgique, et cela depuis plus de vingt-cinq ans, de répondre aux enjeux d'une statistique criminelle intégrée, avec plus ou moins de succès. **Benjamin Mine**, également chercheur à l'INCC, rentre ensuite dans le débat technique en étudiant deux problèmes majeurs qui entravent les possibilités d'articulation des bases de données du système d'administration de la justice pénale: l'absence d'identifiant unique et l'absence d'harmonisation des nomenclatures d'infractions. **Isabelle Detry**, toujours du même institut, propose quant à elle une analyse socio-institutionnelle de l'informatisation des services et de son impact sur l'existence et la qualité des données statistiques, en étudiant plus précisément

⁴ Colloque international, *La production et la diffusion des statistiques pénales en mutation*, organisé par l'Institut National de Criminalistique et de Criminologie (INCC), le Centre d'Etudes du droit et de la justice (CEDJ, UCL), la Cellule interdisciplinaire de Technology assessment (CITA, FUNDP), Maison des parlementaires, Bruxelles, 6 mai 2009.

le domaine de la protection de la jeunesse. Vient enfin, pour clôturer l'état de la situation en Belgique, une contribution collective émanant du milieu universitaire et rédigée par **Julie Louette** (Université catholique de Louvain), **Anne-France Brogneaux** (Facultés universitaires de Namur), **Virginie Detienne** (Agence de stimulation économique) et **Frédéric Vesentini** (Centre de recherche sociologique sur le droit et les institutions pénales). Cet article aborde cette fois la diffusion des données et présente les résultats de projets de recherche fédéraux visant à construire une application dédiée à la diffusion critique des données statistiques: Quetelet.net.

Les deux articles suivants s'offrent ensuite comme des contre-points en présentant ce qui se fait à l'étranger, dans deux pays réputés pour la qualité de leur production statistique: les Pays-Bas et la Suisse. Le texte de **Sunil Choenni** et de **Ronald Meijer** du Centre de recherche et de documentation scientifique du ministère de la Justice (WODC) expose comment aux Pays-Bas, en utilisant un datawarehouse centré sur l'auteur, il est techniquement possible de travailler sur différentes bases de données afin d'obtenir des informations sur l'ensemble de la chaîne pénale. **Steve Vaucher** et **Daniel Fink** de l'Office fédérale de la statistique (OFS) nous emmènent ensuite dans l'univers institutionnel fédéral suisse en évoquant à la fois la question des enjeux de la production et de la diffusion des données.

Les mots de la fin sont laissés à **Bruno Aubusson**, directeur de recherches du CNRS au Centre de recherche sociologique sur le droit et les institutions pénales, qui lors du colloque de mai 2009 a, en guise de conclusions, resitué le débat plus largement en réintroduisant une perspective de plus longue durée et, au passage, en évoquant la situation française.