

French official statistician and ethics: from law to practice

Mylène CHALEIX (INSEE, France) - e-mail : mylene.chaleix@insee.fr

Olivier LEFEBVRE (INSEE, France) - e-mail : olivier.lefebvre@insee.fr

Abstract

Created in 1946, INSEE's missions are to develop and disseminate official statistics to inform economic and social debate in the service of democracy. The 1951 Statistical Act on statistical confidentiality, coordination and obligation is the cornerstone of its mandate and embodies the values of French statistical service. In particular, from the outset it has struck a balance between the conditions under which data is collected and the ways in which it is protected.

However, in addition to being enshrined in law or in initial training programmes, ethics must also be part of everyone's day-to-day practices. In this respect, the organisation adopted for French official statistics is to give statisticians end-to-end responsibility for their process, which includes the production of the final result, but also respect for the values of official statistics. From the design stage to the dissemination of data and studies, the statistician will meet people on numerous occasions who will (re)question the measures taken to contribute to this:

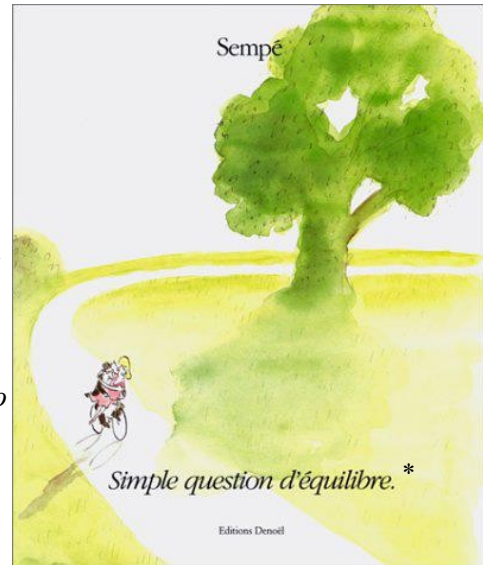
- *the advisability of launching a statistical operation, access to data (existing data or setting up a survey),*
- *compliance with European statistical best practices,*
- *GDPR declarations for individual data,*
- *conditions of access for researchers to the data produced,*
- *treatment of statistical confidentiality in dissemination,*
- *as well as the IT axis (project management, certification, securing access and workstations).*

These different stages provide collective insurance that the ethics of public statisticians are taken into account as closely as possible to the work, while at the same time giving meaning to these issues. Combined with a policy of staff mobility between the various statistical services, they strengthen the ownership by everyone and the cross-fertilisation of approaches, bringing greater security to the whole.

Ethics and practice, the practice of ethics, the example of INSEE

The practice of ethics at INSEE and in the French official statistical service is the result of a balance between ideals and realities, between procedures and their adaptation to circumstances, between theory and practice. Like a bicycle, this balance is only viable if it is accompanied by movement, in other words if it is practised on a daily basis and if its components are continuously adapted.

All official statisticians can therefore rely on a legal framework that gives them a sense of responsibility and protection, as well as on a body of professional values and best practices that are regularly questioned and confirmed in real-life situations. Procedures govern their implementation, while remaining adaptable enough to take into account exceptional circumstances. The way INSEE is organised allows for regular exchanges between colleagues, enabling them to find answers to their questions, and to find the right balance for each of these values.



* Simple question of balance

The legal and organisational framework at INSEE, a vehicle for the values of official statistics

The French National Institute for Statistics and Economic Studies - INSEE - was created by the Finance Act of 27 April 1946, taking over a public statistics activity that had been carried out continuously since 1833. Under the French Statistics Act of 1951, its mission is to collect, analyse and disseminate information about the French economy and society throughout the country. It coordinates the French official statistical service.

In comparison with most other national statistical institutes, INSEE has two important specific features: it carries out economic and social studies based on the data it produces, as well as short-term economic forecasts; it manages inter-administrative registers of people and economic entities on behalf of all stakeholders. A third distinctive feature of INSEE is the way in which its staff are trained: most of the staff recruited undergo initial training in economics and statistics at specialised schools. These schools train both INSEE civil servants and future statisticians for the private sector.

A national and European legal environment that is evolving to adapt to the context and challenges, without losing sight of the fundamentals (providing a framework for data collection, guaranteeing its relevance and data protection).

The challenge: guaranteeing a balance between data collection, use and protection

INSEE operates within a national legal framework that is both relatively old, and therefore firmly rooted in custom and practice, and evolving to take into account changes in context or needs. The founding legislation for this statistical activity dates back to 1951. Initially focused on surveys, it defines the rules of the game in terms of: first data collection (including the obligation to answer), second opportunity (in relation to the need to shed light on a given phenomenon but also the absence of other sources) and third the protection of the data collected (statistical confidentiality). It has gradually been enriched to enshrine professional independence in law and to take into account new data collection methods (use of administrative data and, more recently, data held by private bodies), while at the same time implementing the rules to assess the appropriateness of such data collection and the protection of the data collected.

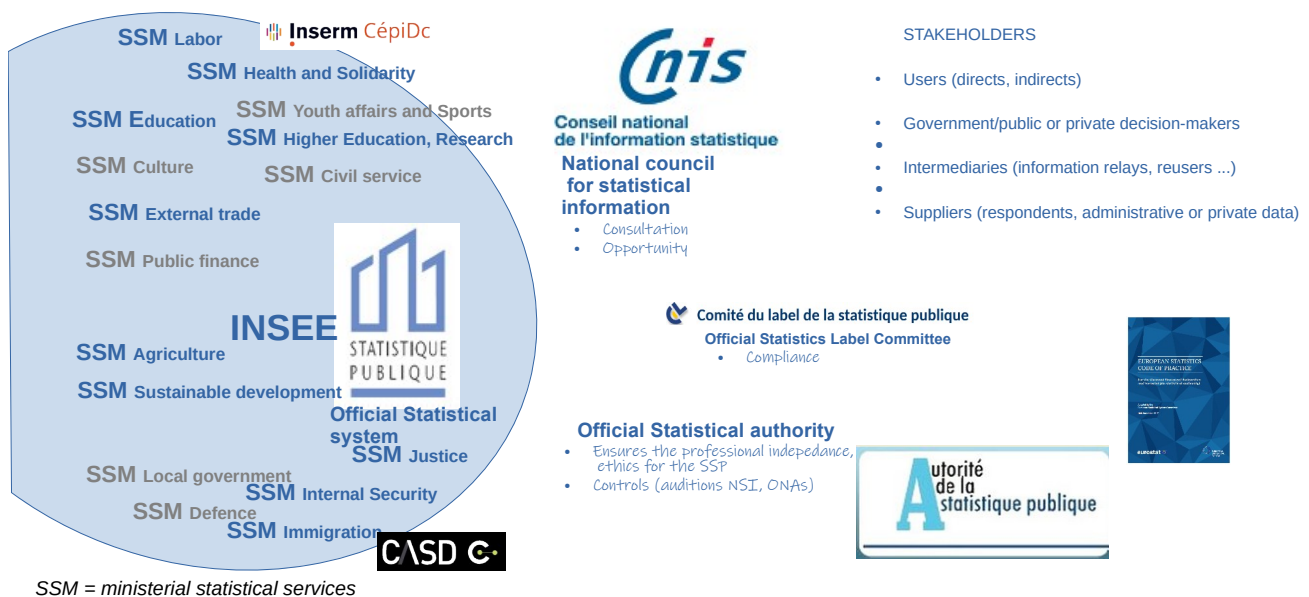
The processing of personal data is carried out in compliance with the 1978 *Data Protection Act*, which has also evolved to adapt to the *General Data Protection Regulation (RGPD)* and to new opportunities for the secure

linking of personal data. The *National Commission for Information Technology and Civil Liberties* (CNIL) is responsible for informing and monitoring its application.

The organisation of official statistics in France is based on 3 pillars.

The 1951 Statistical Act on the obligation, coordination and confidentiality of statistics, revised in particular by the law on the modernisation of the economy of 4 August 2008, created various bodies which play an essential role in the operation and regulation of official statistical activities:

- the official statistical service (SSP), comprising INSEE and 16 ministerial statistical offices (SSM), which carry out statistical operations in their areas of competence;
- the *National Council for Statistical Information* (CNIS), a consultative body for producers and users of official statistics, responsible for monitoring statistical work; its remit is similar to that of ESAC for European statistics, although it is organised differently,
- the *Official Statistics Authority* (ASP), which is specifically responsible for ensuring the professional independence of official statisticians. It is the French equivalent of the ESGAB.



INSEE is responsible for coordinating the methods, resources and statistical work of the SSP and for unifying statistical classifications. The way in which the SSP carries out its activities is supervised by the Official Statistics Authority. The reference framework is defined in particular by the *European Statistics Code of Practice* (CoP).

Also created by the 1951 Statistical Act, the *Official Statistics Label Committee* and the *Litigation Committee for Compulsory Statistical Surveys* deal respectively with surveys on quality and compliance with the obligation to answer. The Label Committee also intervenes, at the request of the ASP, on the conformity of statistics produced by bodies other than those of the SSP (with levels of conformity depending on the levels of labelling requested).

| A strong commitment to sharing data while respecting statistical confidentiality

Unlike a number of NSIs, INSEE has not been entrusted with the task of 'data stewardship' at national level, nor does it claim to do. It does, however, carry out cross-cutting missions in terms of data administration and sharing, for very specific audiences, types of data and domains. The first mission is to share statistical data within the official statistical service (with, in particular, the introduction of a non-meaningful individual identifier internal to official statistics), or with researchers, under conditions defined by a specific body called the *Statistical Confidentiality Committee*. The second is to manage a set of classifications, including geographical ones, and common concepts designed to make it easier to share these data. A third mission, which could be described as 'statistical stewardship', consists in making the data it produces widely available in open data format. Finally, through its management of inter-administrative registers, INSEE carries out an "ID

stewardship" mission designed to facilitate the semantic interoperability of administrative files by means of unique and shared identifiers.

The practice of ethics on a day-to-day basis at INSEE is the result of several factors

A body of good practice and essential values


The challenge: guaranteeing quality, responsiveness and practicality

This legal and institutional framework completes an other one, based on professional values and good practice. In some respects, the former has translated the latter into law, by "making these values enforceable", when necessary (in case of new questions, new opportunities, new challenges). In some others respects, it can be said that these are coherent formalizations, and that the principles and values allow each statistician to "preserve meaning".

One example is the European Statistics Code of Practice, whose 16 principles guide a number of our projects, like the core values adopted by UNECE. This corpus is projected into the daily lives of statisticians through illustrations, behaviours and shared points of attention. The aim is to maintain a high level of quality, in line with our core values, while remaining responsive to real-life situations and all their unexpected aspects.

The six core values adopted by UNECE en 2023

Relevant	Impartial	Transparent
Professionally independent	Protects privacy	Collaborative



Guides, procedures, audits and action plans to help with implementation

The challenge: to remain solid and not lose our sense of purpose

Beyond these principles and values, official statisticians rely on three pillars. The first concerns shared methods for their actions, to ensure that they are correctly implemented. The second involves procedures for day-to-day operations and also action plans to improve or implement investments. The third involves audits designed to provide an external view of our practices. These audits allow to assess the effectiveness (are they achieving the desired goal?) and the efficiency (with the adequate allocation of resources?). The challenge is for all statisticians to have concrete, shared benchmarks for the implementation of their work, without being locked into total formatted processes, and to be able to evaluate and develop them.

An organisation that gives responsibility to the project designers and to each of the players involved, and enables them to look at each phase of the process from different angles.

The challenge: integrating diverse contributions without diluting responsibilities

INSEE has adopted a work organisation that gives full responsibility to the designer of a statistical operation (from data collection to analysis and the provision of initial results). This organisation provides the technical skills and the outside expertise needed to successfully complete the operation, and to continuously improve it especially in case of a recurrent operation. The manager in charge of the process surrounds him/herself first with experts in the field to target the collection of information according to the stakeholders' expectations.

After, he/she works with experts in methodology or survey design¹, as well as with architects and IT developers for high-performance and secure implementation of the collection and processing. He/she relies on interviewers, then controllers, for these collection and processing operations... Checkpoints (opportunity opinions, statistical compliance opinions, IT security approvals) allow outside parties to express their views, so as to improve the system or, if not, to express their opposition to its implementation.

A practice that enables us to adapt to external shocks or particular contexts (passionate subjects, sensitive issues, public scrutiny) without sacrificing our values.

The challenge: knowing how to relax the rule when necessary without losing its spirit, or knowing how to go further than the proposed framework.

The procedures are clear, just restrictive enough to guarantee their efficiency (by preventing everyone from re-inventing them) and to ensure the necessary transparency, while retaining the possibility of adapting them to the circumstances. For example, during the Covid crisis, official statistics were able to carry out a survey on companies and published the results less than 30 days after the request was made. Accelerated procedures to examine the appropriateness and conformity of statistics, combined with a flexible and secure IT system, enabled a crucial question to be answered quickly (how did businesses adapt to the initial containment?) without sacrificing the quality or security of data collection and processing.

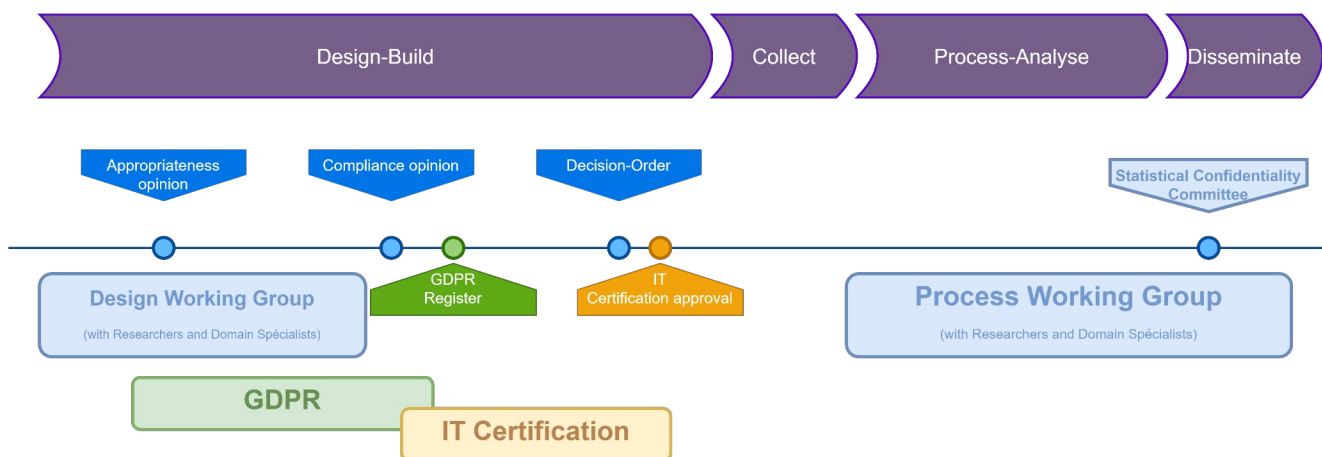
A shared mindset based on the values of quality, confidentiality and independence

The challenge: a common compass for day-to-day use

For all of this to work optimally, it is essential that everyone shares a vision of the aims of the statistical operations to which they contribute, but also of the values that underpin its implementation, and of their own role in carrying them out.

The values are part of the teaching received in the schools, and then they are continuously infused through exchanges between colleagues, additional training, internal communication and managerial practices. One of the challenges is to link everyday actions and routine or exceptional work to these values, so as to embody them. The level of formalisation of the values has so far remained minimalist, but it may be necessary in the future to expand them to meet the growing expectations of users and facilitate communication on these values.

In practice, this is illustrated by the classic procedure for an official statistics survey



Any official statistics survey project must be the subject of an opinion from the CNIS, which checks the general interest and usefulness of the statistical operations presented in the work programmes. This opinion, issued by the president of the thematic commission responsible for the survey, ensures that it meets a public interest need

¹questioning, questioning protocol, whether or not to use multi-mode, sampling

(principle of relevance) and that it does not duplicate other sources already available - statistical or administrative survey, management file, etc. (principle of minimising the burden on respondents).

In order to be carried out within the framework of the 1951 Statistical Act, any statistical survey project presented by a public statistics producing service must obtain a label of general interest and statistical quality. After receiving an opinion, the person in charge of the survey (within the producing service) prepares a file (which, since 2024, includes the main headings of the SIMS (Single Integrated Metadata Structure) standard for user-oriented quality reports, supplemented by methodological appendices). Next, the Official Statistics Label Committee gives its opinion which then ensures that survey meets statistical quality criteria. These criteria concern the consultation with users, the collecting and sampling methods², the relevance of the questioning and the adaptation of the dissemination to the stated objectives. The file also includes the results of questionnaire testing. It also ensures that the survey does not place an excessive burden on respondents, that consultation has taken place with the stakeholders and that the wishes expressed by the CNIS during the debate on the appropriateness have been taken into account.

At the design stage, the survey manager can call on experts from various cross-functional units:

- for the design of their questionnaire (Statistical Methods Unit) and specialists in the field (internal or external to the SSP) ;
- for sampling (Statistical Methods Unit);
- on the organisation of its survey (Survey Project Management Units);
- on the definition of dissemination products (Dissemination and Regional Action Directorate) and the treatment of confidentiality (Statistical Methods Unit)...

Illustration with TEO: Trajectories and Origins survey (2008 - 2019 editions)

The Trajectories and Origins survey (*Trajectoires et Origines* - TeO) attempts to measure the impact of origins on access to the main resources that define one's place in society: housing, education, language skills, employment, public services and social benefits, health, social and family relations, nationality, citizenship, etc.). It looks at the relationship between origins and other categories of distinction in French society (gender, class, age, neighbourhood, etc.) in order to analyse the processes of integration, discrimination and identity-building within French society as a whole. The second edition of the survey (TeO2), was carried out in 2019-2020, updated the results of the first survey (TeO1) on these different themes ten years later.

Following strong public demand for information on immigrants and their descendants, and their integration (lack of information in administrative data or in the census), INSEE and INED - the French Institute for Demographic Studies, a public research body) have joined forces to offer an original survey on trajectories and origins. The special feature of the survey is that it deals with issues classified as sensitive under the Data Protection Act. These include in particular issues relating to racial or ethnic origins, religion, health or civic life (trade union or political opinions).

For the first edition, a multidisciplinary working group (statisticians, demographers and researchers) was set up to design the questionnaire between 2006-2007 (with a focus group and field test), the associations concerned were consulted, and the CNIS held discussions on the sensitivity of the data (in May and October 2007) in addition to the usual opportunity and compliance stages. The opinion of the CNIL was also sought out.

Driven once again by public demand, a new edition was included in the survey programme. During the preparation of TeO2 (2016-2019), the same stages were repeated (design group, opinion of opportunity; compliance, opinion requested from the CNIL and registration with the GDPR and their recommendations were included in the protocol. The recommendations were on the experimental nature of the survey on the 3rd

² sampling plan, data adjustment method, treatment of non-responses to guarantee the reliability of the results, etc.

generation or the possibility of not responding on sensitive topics.

In addition to surveys, the data-producing services also ask the CNIS to access administrative data (1951 Statistical Act - Article 7bis) or private data (1951 Statistical Act - Article 3bis) and to obtain an opinion. The question of proportionality, which is not addressed through the prism of the response burden, requires particular attention. In France, this attention is ensured by the Data Protection Act, which states that "data collected for a given purpose must remain adequate, relevant and not excessive, and that the list of data must be limited to what is strictly necessary to meet the stated purpose". With the development of statistical identifiers and the forthcoming introduction of dedicated applications, consideration is being given to the introduction of a linking procedure in order to control the objectives and, once again, limit the information linked to that which is necessary.

Other safeguards for data confidentiality and security

DPIA: a procedure to protect personal data

The General Data Protection Regulation (GDPR), transposed in 2018 into French law by the changes made to the Data Protection Act, imposes obligations on all data providers. These obligations essentially revolve around two main principles: a principle of transparency and a principle of controlling and limiting the risks of process' impact on the privacy of data subjects. The principle of transparency requires data subjects to be informed, in particular of the purpose of the processing operation, the data used, how to exercise their rights, and how to respond to their requests, as well as having access to the documentation on the processing operations (see letter of notification for respondents and information on insee.fr). The data provider must also ensure that the processing carried out complies with the principles of necessity, minimisation and proportionality with regard to the data processed.

The French version of the GDPR requires all processing of personal data to be recorded in a register kept by the service responsible for the processing, as well as realising an impact assessment on the consequences for individuals in the event of a proven risk. The register of processing operations describes the purposes and objectives of each personal data processing operation, the categories of data used, the categories of data subjects, the parties involved (producers, recipients, processors) and the storage periods. The register is used to compile data processing description sheets, which are published online at insee.fr or communicated directly to respondents.

The Data Protection Impact Assessment (DPIA) is a tool that can be used to design a processing operation that complies with the GDPR and respects privacy. It concerns the processing of personal data that is likely to give rise to a high risk for the rights and freedoms of data subjects.

To carry out a DPIA, it is necessary to:

1. define and describe the context of the processing operation(s) in question;
2. analyse the measures guaranteeing compliance with the fundamental principles of proportionality and necessity of the processing operation, and protection of the rights of the data subjects;
3. assess the privacy risks associated with data security and ensure that they are adequately addressed; formally validate the DPIA in the light of the above elements, or decide to revise the previous stages.



Source : CNIL

The data protection impact assessment includes a more detailed description of the processing operation than that contained in the register of processing operations. It also deals with the factors that will enable the sensitivity of the processing to be assessed, in particular with regard to its scale (the entire population or a

sample?), the variables processed (in particular with regard to their sensitivity within the meaning of the GDPR) and the nature of the processing itself (does it involve cross-referencing several files?). In particular, it includes an analysis of the risks in terms of data security or breach of privacy, together with remedial measures. This document is drawn up by a specialised unit at INSEE, in conjunction with the designer of the processing operation, who thus benefits from essential expertise in analysing these issues.

|... and another procedure for data integrity and security

Securing data at every stage of the process is also a concern for public statisticians, and one that is growing with the rise in cyberattack. At INSEE, information systems are subject to security certification, involving experts from the Information Systems Directorate and external experts. Once again, the process will be scrutinised, but this time from the point of view of IT security: availability, integrity, confidentiality and traceability. As far as surveys are concerned, since the data collection tools are pooled within a survey network, the certification of a survey is based on what has been done for the network.

But being able to think outside the box when necessary

*|How can we ensure that a sensitive project is acceptable?
The challenge: legitimacy beyond the ability to do things*

In 2021, INSEE has launched an ambitious project called RÉSIL, which aims to build a statistical register of individuals and housing based on the linkage of various administrative data.

In this project, INSEE is convinced that the legal and ethical issues are just as important as the technical challenges and require special attention, with the support of outside experts. In addition to the technical and legal capacity to make RÉSIL work, we also needed to acquire legitimacy, through a "social mandate". An ambitious process of consultation with civil society was therefore carried out on this project, in parallel with the statistical engineering work. In particular, it relied on a working group from outside official statistics, which operated very intensively for 6 months, from May to October 2022. This was an extra-ordinary approach, going beyond the usual consultation processes, but adapted to the very specific challenges of this project.

The group was not opposed to the principle of RÉSIL, as an exclusively statistical register fed by various sources and allowing record linkages, but had confidence in the institution from a technical and ethical point of view to build and operate it in accordance with the principles of good practices.

However, the group considered that given the nature of RÉSIL and the current context regarding the use of personal data, marked both by greater circulation of this data and increased vigilance over its use, it is necessary to set and make visible the rules of the game, to rely on outside views so that various national authorities or agencies can guarantee them and set the limits. They also felt that it was essential to continue the transparency and consultation efforts undertaken by INSEE over the long term.

This experience of consultation was very challenging, but very useful for INSEE, as it enabled us to identify certain concerns that might emerge about such a system, certain challenges in terms of communication and consultation on RÉSIL and the use of external data which in turn helped us to improve the design of our project.

*|Sensitive data
The challenge: proportionate data collection*

Knowing not to go too far for reasons linked to social acceptability: in the example of TeO, the exceptions provided for in the regulations on sensitive data (on racial or ethnic origins in particular) allowed INSEE to collect information on skin colour with a consensus reached on self-perception which included a free response without predefined categories. However, the principle of collecting this information was debated and, noting that social acceptance of this type of questioning was not guaranteed, INSEE and INED decided to abandon it.

Another example with RÉFIL is that some of the administrative sources envisaged were not retained following discussions in the consultation group, which warned of the risks (proportionality and acceptability). In this respect, benefiting from an outside perspective is essential in assessing the principle of proportionality.

And communicating our values

As part of RÉFIL, in addition to the transparency and collaboration with stakeholders practised during the consultation group, it was decided to communicate more widely on the values underlying the construction of the directories, also by highlighting professional independence, confidentiality, relevance and impartiality.

Downstream: more opportunities to ask questions

Once the data has been collected and processed, there are other opportunities for the statistician to question the relevance of the results and the associated documentation. For some surveys in particular, an analysis working group has been set up, bringing together researchers specialising in the field as well as official statisticians. This working group provides both an outside view of the data and exchanges on the initial results, or on the content and documentation associated with the files disseminated (FPR - Production and Research Files, see below). This working group is generally made up of people who are already members of the consultation working group set up when the survey was designed or redesigned.

New challenges arise when disseminating results: taking into account statistical confidentiality when disseminating data

The challenge: striking a balance between open data and data protection

To do this, the survey manager can call on the experts in the Geographical Methods and Reference Frameworks Unit to ensure the statistical confidentiality of the tables produced, which is particularly important for business surveys. For the dissemination of localised data in particular, they can also call on the skills of the experts in the Dissemination Unit.

The 1951 Statistical Act on the obligation, coordination and confidentiality of statistics set up a *Statistical Confidentiality Committee* to ensure compliance with statistical confidentiality rules. This committee issues an opinion on requests to disclose data covered by statistical confidentiality for the purposes of official statistics or scientific or historical researches. The data concerned may have been collected as a statistical survey or transmitted by administrations or private operators to the official statistical service in order to compile statistics. The applicant submits a file explaining the purpose of the study and justifying the list of data required. And after a favourable opinion, access is generally granted at the CASD (secure access centre), which allows the applicant to work on the data and extract results that respect statistical confidentiality.

A simplified procedure has been put in place for researchers with production and research files (files less rich than those available at the CASD to limit the risks of re-identification).

The importance of a shared culture among INSEE staff

INSEE staff undergo a high-level initial training course specialising in statistics and economics at one of the two schools of the *Groupe des écoles nationales d'économie et statistique* (GENES) or at the INSEE Training Center in Libourne (CEFIL). A continuous training programme ensures that staff develop their skills throughout their career. Training is not the only means of ensuring that the ethical values of official statistics are embraced. The mobility policy enables INSEE staff to regularly broaden their field of expertise, either by moving into different professions within the same field, or by moving between different fields of activity. In this way, they can start their career in a position of methodological expertise before taking responsibility for a survey, or vice versa. They are also encouraged to work in the ministerial statistical offices and to help disseminate good ethical practice throughout the official statistical service.

Conclusion

The environment in which we carry out our work is constantly changing. It offers new opportunities in the way we access and process data; it places us in a competitive situation with data producers who are now able to produce their own analyses or assessments; it confronts us with new needs for data or analyses to understand demographic, economic or social phenomena, it requires also quicker answers.

In the face of competition, the demands for relevance and quality are essential, as are strict compliance with data protection and transparency in data processing; they must be accompanied by a 'know-how' that enhances the value of the data itself. We need to learn to communicate our values better.

New opportunities give us new responsibilities: we must make the best use of them, while respecting the framework of necessity and proportionality and ensuring transparency about their use. This is a guarantee of the trust that citizens and users of our data place in us.

Being responsive and adaptable without abandoning our values or the quality of our output, implies a number of requirements: technicality, expertise, quality of tools for collecting, processing and making data available, the existence of tried and tested processes, but also a shared vision of the purposes of our work, enabling us to optimise its implementation, in a secure manner, depending on the circumstances and in line with our values. It's a question of knowing how to make "tailor-made" products, making optimum use of the "machine tools", without the process becoming an assembly line.

Tailor-made vs Taylor-made...

Legislative framework

[Loi n°51-711 du 7 juin 1951](#) sur l'obligation, la coordination et le secret statistique (Statistical Act)

Décrets [n°2009-250 du 3 mars 2009](#) modifié et [n°2009-318 du 20 mars 2009](#) pour les missions de l'ASP, le Cnis et les comités du label de la statistique publique, du secret statistique et du contentieux des enquêtes statistiques obligatoires (ASP and CNIS Decrees)

[Loi n°78-17 du 16 janvier 1978](#) relative à l'informatique, les fichiers et aux libertés (Data protection Act)

[Règlement \(UE\) n° 223/2009](#) révisé du Parlement européen et du Conseil du 11 mars 2009 relatif à la statistique européenne (European Statistical Act)

[The French official statistical system](#) on the website insee.fr (english version)

Bibliographies

- Anxionnaz I. et Maurel F., [The National Council for Statistical Information](#): The quality of Official Statistics also depends on consultation, in *Courrier des statistiques* n° 6 (july 2021)
- Roth N. et Christine M., [The Label Committee](#): A Governing Body Ensuring the Quality of Official Statistics, in *Courrier des statistiques* n° 5 (december 2020)
- Bureau D., [Ensuring Independent Quality Statistics](#): The French Official Statistics Authority Ten Years After, in *Courrier des statistiques* n° 5 (december 2020)
- Redor P., [Confidentialité des données statistiques](#) : un enjeu majeur pour le service statistique public, in *Courrier des statistiques* n° 9 (juin 2023) – in French only
- Tavernier J.-L., [Un système statistique intégré à l'administration centrale](#), in *Courrier des statistiques* n° 1 (décembre 2018) – in French only
- Isnard M., [Qu'entend-on par statistique\(s\) publique\(s\) ?](#), in *Courrier des statistiques* n° 1 (décembre 2018) – in French only
- Lefebvre O. [Towards a "social mandate" for the French project of a statistical directory of individuals and dwellings by France](#) - Unece juin (2023)