LA MISE EN MUSIQUE D'ENQUÊTES MULTIMODES

Eric Sigaud et Benoît Werquin***

Pour l'Insee, le développement d'enquêtes multimodes relève d'une orchestration délicate et rigoureuse. Il s'agit, entre autres, de construire des questionnaires internet cohérents avec des questionnaires enquêteurs, de coordonner le travail des enquêteurs et gestionnaires mobilisés sur des collectes multiples dans différents modes, le tout en conservant une mise en œuvre harmonieuse et efficace prenant en considération les contraintes de ressources et de calendrier. Ces enjeux de construction et de transformation sont autant de tonalités de la démarche mise en œuvre par l'institut depuis une dizaine d'années. Celle-ci est rythmée par des implémentations en vraie grandeur sur des opérations d'enquêtes, d'abord auprès des entreprises puis des ménages, d'abord** via **internet puis avec les enquêteurs, d'abord sur des protocoles simples ne mobilisant qu'un seul mode de collecte puis sur des protocoles complexes, avec plusieurs collectes, dans des modes différents, simultanément.

Depuis le lancement de cette évolution structurante pour la statistique publique, le même schéma s'applique : penser d'abord le métier, l'exprimer, le conceptualiser, puis passer aux phases de mise en œuvre ou de passage à l'échelle. L'expérience acquise à chaque étape favorise une construction progressive de l'outillage conceptuel et technique des enquêtes « réellement » multimodes, dans toute leur complexité.

For INSEE, the development of multi-mode surveys is a delicate and rigorous orchestration. Among other things, it is a matter of constructing Internet questionnaires that are consistent with survey questionnaires, and coordinating the work of interviewers and managers mobilised for multiple collections in different modes, all while maintaining a harmonious and efficient implementation that takes into consideration resource and schedule constraints. These issues of construction and transformation are all part of the approach implemented by the Institute over the last ten years. This has been punctuated by full-scale implementations of survey operations, first with companies and then with households, first via the Internet and then with interviewers, first with simple protocols using a single collection mode and then with complex protocols, with several collections, in different modes, simultaneously.

Since the launch of this structuring evolution for official statistics, the same pattern has been applied: first think about the business, express it, conceptualise it, then move on to the implementation or scaling-up phases. The experience acquired at each stage favours the gradual construction of the conceptual and technical tools of "truly" multimode surveys, in all their complexity.

* Maître d'ouvrage délégué du programme Métallica, département des Prix à la consommation et des enquêtes ménages, Insee, *eric.sigaud@insee.fr*

** Chef de projet informatique de Métallica, service national du développement informatique, Insee Hauts de France, *benoit.werquin@insee.fr*

Pour nombre d'enquêtes de la statistique publique française, auprès des entreprises ou des ménages, la collecte par internet constitue une alternative à la collecte papier depuis quelques années déjà[1]. Elle devient également une alternative ou un complément pour la collecte en face-à-face ou par téléphone, créant, de fait, de nouveaux protocoles de collecte, plus complexes, qualifiés de « multimode ».

Au-delà des délicates questions statistiques ou méthodologiques posées par l'avènement de tels protocoles[2], il convient de se pencher, à la lumière de l'expérience acquise, sur la complexité opérationnelle induite par le multimode.

Comment construire un ensemble d'outils cohérents capable de rendre les services attendus ? Quels sont précisément les services attendus ? Comment les organiser ? Sur quel acteur repose quelle tâche ? Quelle offre de services mettre en œuvre ? Dans quel ordre ?

La démarche s'apparente en quelque sorte à l'écriture par l'Insee de sa partition et de ses arrangements pour arriver à une collecte multimode harmonieuse[3]. Des premiers solos, opérations mono-modes, ont d'abord permis à l'institut de rôder les concepts, les outils et les processus. Puis des déclinaisons « polyphoniques », où cette fois différents modes sont mis en œuvre indépendamment les uns des autres, ont permis de caractériser chaque mode de collecte, constituant une étape préalable qualifiée de « poly-mode ». Elle prépare à la mise en œuvre de protocoles multimodes à proprement parler, où les différents modes de collecte doivent cette fois agir de concert.

POUR PASSER D'UN SEUL MODE DE COLLECTE À PLUSIEURS

Le développement de la collecte par internet, en complément d'un autre mode, s'accompagne immanquablement de son lot de questions pratiques : comment construire un questionnaire *web* ? Doit-on avoir le même questionnement dans les différents modes ? Comment contacte-t-on les répondants ? Comment les relance-t-on ?

> *« Conceptualiser ce qui est déjà fait et ce qu'il y aurait à faire. »*

Si le multimode pose des questions complexes, sur la qualité et l'exploitabilité des données collectées, il faut auparavant résoudre la question de la faisabilité : comment fait-on pour conduire une collecte recourant à plusieurs modes, dont un nouveau ? Un premier réflexe est de s'appuyer sur les pratiques courantes pour définir les nouvelles, pour tenter de reproduire ce qui est déjà fait. Cependant, la complexité induite par le nouveau mode de collecte (développement d'un nouveau questionnaire, nécessité de gérer des identifiants/mots de passe, etc.) et ses particularités (pas toujours de phase d'approche par l'enquêteur, contexte de réponse moins contraint, etc.) conduisent à pousser la réflexion un cran plus loin : conceptualiser ce qui est déjà fait et ce qu'il y aurait à faire.

1. Au tout début des années deux-mille, le service statistique du ministère de l'Industrie lançait les premières enquêtes par internet de la statistique publique. En « régime » courant en 2021, 40 enquêtes auprès des entreprises utilisent le mode *web*, et 4 enquêtes auprès des ménages, dont l'enquête Emploi.
2. Voir sur ces questions l'article de François Beck, Laura Castell, Stéphane Legleye et Amandine Schreiber sur « *Le multimode dans les enquêtes auprès des ménages : une collecte modernisée, un processus complexifié* » dans ce même numéro.
3. [NDLR] Les métaphores musicales ne surprendront pas les lecteurs des numéros précédents sur les sujets connexes, comme (Cotton et Dubois, 2019 ; Haag et Husseini-Skalitz, 2019 ; Koumarianos et Sigaud, 2019).

LE «POLY-MODE» (PREMIÈRES VOCALISES)

La première étape de conceptualisation enclenchée par l'Insee, de loin la plus mûre aujourd'hui puisqu'en place depuis plusieurs années[4], est la **génération automatique d'instruments de collecte** (appelée parfois par abus de langage génération de questionnaires). Elle repose sur le principe des **métadonnées actives** (Bonnans, 2019) et part de l'idée simple qu'un processus répété, en l'occurrence le développement d'un questionnaire d'enquête, gagne souvent à être automatisé.

Pour conceptualiser un processus, celui de « fabrication » du questionnaire, le GSBPM[5] est le standard référence en la matière. Il suggère un premier découpage, entre une phase de conception et une phase de construction. Dans la première, on spécifie un questionnement, en dehors d'un mode donné. Dans la seconde, on construit des instruments de collecte[6] dédiés à un mode. À l'Insee, ce découpage structure les outils *Pogues* et *Eno* (Cotton et Dubois, 2019 ; Koumarianos et Sigaud, 2019) : *Pogues* permet la spécification d'un questionnaire dans une application tandis que *Eno* supporte la génération des instruments de collecte ***(figure 1)***.

Cette approche générative permet de capitaliser sur les bonnes pratiques entre questionnaires et de mettre en œuvre les fonctionnalités propres à chaque mode :

- des codes-barres pour les questionnaires papier ;
- une accessibilité et un fonctionnement sur différents supports (*smartphone*, tablette, etc.) pour les questionnaires *web* ;
- une navigation et une ergonomie adaptées pour le questionnaire déroulé par un enquêteur ;
- une contextualisation du questionnement au mode (liste déroulante en *web*/carte code en face-à-face vs champ libre sur papier), etc.

Ainsi, la production de questionnaires peut changer d'échelle, passant du développement *ad hoc* de questionnaires spécifiques à chaque mode et à chaque enquête, à une automatisation de la production d'instruments de collecte, pour plusieurs modes de collecte, s'appuyant sur un même outil de spécification et rationalisant les travaux de développement : formellement, on ne conduit plus de développement *ad hoc*.

La logique « **mono-mode** », où le processus et les outils sont pensés et mis en œuvre pour un mode donné, laisse la place à une logique « **poly-mode** », où processus et outils sont conceptualisés pour supporter plusieurs modes de collecte et plusieurs cas d'enquête (*web*, papier, téléphone, plusieurs séquences de collecte, etc.).

On comprend alors tout l'intérêt de l'approche par les métadonnées actives, qui permet la génération automatique des questionnaires, et plus largement des « instruments de collecte » (Koumarianos et Sigaud, 2019). L'approche ne se limite d'ailleurs pas aux seules phases de conception et de construction des instruments de collecte. Et après ces premières phases du GSBPM[7], c'est le processus de collecte dans son ensemble qui pourra faire l'objet

4. Les premiers travaux de génération de questionnaires à l'Insee remontent à 2013 et portaient sur le questionnaire de l'enquête sectorielle annuelle de la sphère entreprise.
5. *Generic Statistical Business Process*, ou modèle générique de processus de production statistique, mis au point dans le cadre de l'Unece. Voir (Unece, 2019) et (Erikson, 2020) pour un exemple de mise en œuvre opérationnelle.
6. Un instrument de collecte est l'instanciation d'un questionnement dans un contexte opérationnel (questionnaire internet, questionnaire adapté à un appel téléphonique d'un enquêteur, adapté à la visite en face-à-face d'un enquêteur, questionnaire papier).
7. Formellement, le GSBPM distingue 8 phases : la définition des besoins, la conception, la construction, la collecte, le traitement des données, l'analyse, la diffusion et l'évaluation.

Figure 1. Le poly-mode, première étape vers l'utilisation des métadonnées au service de la collecte

*DDI : *Data Documentation Initiative*

Pour plus d'information sur la génération automatique d'instruments de collecte, voir (Cotton et Dubois, 2019 ; Koumarianos et Sigaud, 2019).

d'une conceptualisation afin de permettre son instanciation selon l'enquête, son mode et son protocole. Ainsi, des métadonnées telles que des découpages en opérations (plusieurs séquences, vagues de collecte), des calendriers d'opérations de collecte (dates de début/fin de collecte, de relance), ou des caractéristiques de protocoles (concurrentiel ou séquentiel, avec ou sans phase de repérage, avec ou sans reprise enquêteurs) pourraient être également « activées » et permettre le pilotage des processus mis en œuvre.

Un exemple de la transformation apportée par la conceptualisation illustre parfaitement l'apport de cette approche par les métadonnées : celui du passage de la fiche-adresse (document contenant les coordonnées d'un ménage à enquêter) à l'unité enquêtée ***(encadré 1)***. Le concept d'unité enquêtée désigne les unités de compte d'une opération de collecte, qu'elle soit enquêtée par le *web*, le téléphone, le face-à-face, le papier. La notion permet notamment d'associer des données collectées par un questionnaire et les informations relatives au déroulement de la collecte (du type paradonnées), indépendamment du mode.

CONCEPTUALISER AVANT DE METTRE EN ŒUVRE

Définir précisément le processus à mettre en œuvre, s'appuyer sur le GSBPM pour le découper en « phases », ensemble d'actions métier cohérentes, sont des préalables à toute mise en œuvre dans les différents modes. Une réflexion décorrélée conduirait à construire des outils et processus incohérents les uns avec les autres ou incompatibles avec les besoins les plus exigeants, tels que des collectes *web* et enquêteurs concomitantes et concurrentes, ou encore un suivi du processus consolidé entre les modes.

> *« C'est ce découpage fin du processus qui va permettre une réutilisation et une mutualisation dans le contexte des enquêtes auprès des ménages. »*

Ceci n'empêche pas une approche incrémentale, tant dans les opérations que dans la complexité considérée. Ainsi à l'Insee, les premiers développements ont concerné les enquêtes auprès des entreprises, par nature plus homogènes que les enquêtes auprès des ménages.

LES ENQUÊTES AUPRÈS DES ENTREPRISES (PREMIÈRE SYMPHONIE)

Le projet *Coltrane*[8] de l'Insee a mis en place des services automatisés et partagés à destination des enquêtes auprès des entreprises, afin de permettre une collecte par internet et par papier. *Coltrane* portait un enjeu fort de rationalisation des coûts, notamment sur les travaux de spécification, de recette et de développements informatiques. La plateforme et les services associés instancient un processus, certes adapté aux enquêtes auprès des entreprises, mais s'inscrivant dans un cadre conceptuel plus global. Notamment un découpage fin du processus, s'appuyant sur le GSBPM, permet une séparation des phases d'approche ou d'accès du répondant, des phases de questionnement à proprement parler. Voit alors le jour un portail « Mes enquêtes » ***(figure 1)***, lequel propose un « tableau de bord » pour les personnes répondant en entreprise, interrogées plusieurs fois et souvent dans plusieurs enquêtes en même temps.

8. (Collecte Transversale d'Enquêtes) Aujourd'hui c'est une offre de service complète qui permet aux concepteurs d'enquête : de collecter les réponses par internet, de disposer d'un référentiel de contacts au sein des entreprises et de le gérer, d'envoyer des courriers, des mails et des questionnaires papier aux entreprises qui le souhaitent, le tout en disposant d'un système d'assistance dédié (Haag et Husseini-Skalitz, 2019).

De même sont mis en place une offre de courriers standardisés[9] ainsi que des processus d'envois automatiques et intégrés, pour l'ensemble des enquêtes auprès des entreprises[10].

C'est ce découpage fin du processus, notamment en phases d'accès aux enquêtes et de réponse à un questionnaire, qui va permettre une réutilisation et une mutualisation dans le contexte des enquêtes auprès des ménages.

Encadré 1. Un changement de partition: de la fiche-adresse à l'unité enquêtée

Les enquêtes auprès des ménages de la statistique publique française se sont longtemps appuyées sur une collecte en face-à-face, habitude qui se retrouve dans la dénomination usuelle, à l'Insee, de l'objet « enquêté ménage » sous le vocable de « fiche-adresse ».

Une fiche-adresse, dans le vocabulaire de l'Insee, désigne au départ la fiche papier sur laquelle sont indiquées les coordonnées d'un enquêté, individu, ménage ou logement qu'un enquêteur doit interroger.

Dans un contexte mono-mode, il y a une certaine logique conjointe entre la production de ces documents papier à destination des enquêteurs, le tirage de l'échantillon et la configuration des applications et des flux pour la mise en œuvre d'une collecte. On retrouve, à chaque fois, peu ou prou, les mêmes informations que celles contenues sur ces « fiches-adresses ». Ainsi, par un abus de langage, le répondant est également dénommé « fiche-adresse » et tout le monde se comprend.

Mais que devient cette fiche-adresse dans une collecte *web* ou téléphone quand aucun enquêteur ne se déplace ? Faut-il encore imprimer ce papier ? Que devient ce papier quand on a surtout besoin de coordonnées téléphoniques plutôt que d'une adresse physique ? Comment reporter sur ce papier les informations provenant d'une collecte *web* concomitante ?

Autant de questions qui nécessitent une réflexion préalable sur le processus (approche ou repérage, prise de contact/rendez-vous ou accès au questionnaire en ligne, récupération des données collectées, etc.) avant de pouvoir décliner l'enquête dans un nouveau mode.

La **fiche-adresse** devient alors **unité enquêtée** et celle-ci traverse les différentes phases du processus de collecte, de sa « naissance », dans les berceaux de l'échantillonnage, à son « éducation », qu'elle soit sous la vigilance bienveillante d'un enquêteur ou livrée aux interactions plus « chaotiques » d'une modernité connectée, jusqu'à son « passage à l'âge adulte », transformée par des traitements en variables statistiques. La conception de ce **« cycle de vie de l'unité enquêtée »** permet d'identifier **les concepts et les événements** nécessaires pour retracer l'histoire vécue par l'unité enquêtée. Ensuite on pourra identifier des objets qui peuvent être partagés au sein des différents systèmes : coordonnées, qualification du type de non-réponse, etc. Ces précisions sur les événements peuvent relever d'un mode en particulier – essai de contact ou de repérage pour les enquêtes mobilisant des enquêteurs téléphoniques ou se déplaçant en face-à-face, informations d'authentification pour les enquêtes mobilisant le *web*, notices papier pour celles offrant la possibilité de répondre par papier, etc. – ou se décliner dans chaque mode – non-répondant, date de début/fin de collecte, relance, etc.

9. Ces modèles de courriers mutualisés sont par ailleurs conformes aux préconisations du comité du Label de la statistique publique.
10. Avec les services spécialisés dans la production éditique.

LES ENQUÊTES *WEB* AUPRÈS DES MÉNAGES (VARIATION SUR LE MÊME THÈME)

Si les premières opérations ont concerné les enquêtes auprès des entreprises, l'Insee a depuis étendu la réflexion et les travaux aux enquêtes auprès des ménages.

Pour envisager de réutiliser dans un contexte « ménage » les services mis en œuvre dans un contexte « entreprise », même par « morceaux », c'est l'analyse fine du processus de collecte qui permet de proposer des scénarios. Les services rendus et les outils mis en œuvre, correctement découpés, peuvent être ensuite « recomposés ». En l'espèce, il a fallu opérer quelques variations par rapport à l'expérience *Coltrane*.

Le questionnement statistique a d'emblée été analysé comme un même besoin métier : il pouvait donc être mis en œuvre par les mêmes services techniques, moyennant un paramétrage pour offrir une contextualisation adaptée à chaque type de répondant ; cela concernait :

- les premières pages du questionnaire (contenant notamment le cadre juridique de l'enquête) ;
- la charte graphique du questionnaire (logo des services statistiques partenaires par exemple) ;
- les mécanismes d'authentification sur le site *web* (un thème et des consignes différentes) ;
- les formules d'assistance aux répondants.

En revanche, pour l'accès au questionnaire, les besoins divergent sensiblement entre les deux univers. Les répondants en entreprise sont en effet plus familiers des enquêtes de la statistique publique, car ils sont souvent interrogés par plusieurs d'entre elles. Dans un contexte « entreprise », un portail d'accès doit donc répondre à des besoins de visibilité sur l'ensemble des enquêtes pour lesquelles un répondant est sollicité ; il doit également prendre en compte des liens parfois complexes entre la personne « contact » et les entreprises pour lesquelles il est autorisé à répondre (par exemple, les comptables sont souvent répondants pour plusieurs entreprises différentes pour plusieurs enquêtes différentes). Ainsi, le choix a porté sur un portail « *Mes enquêtes* », un site où le répondant peut accéder aux différents questionnaires qui le concernent, avec des rappels sur les échéances. Sont également implémentées des fonctions de mise à jour des coordonnées personnelles, informations particulièrement importantes dans un contexte où les ré-interrogations d'une même entreprise sont nombreuses (à chaque vague de certaines enquêtes, ou par différentes enquêtes).

Le besoin des ménages est tout autre : les répondants ne sont en général sollicités que pour une seule enquête[11], ils ne sont pas forcément familiers des concepts d'enquête de statistique publique et sont souvent moins enclins à y consacrer du temps. Dans ce contexte « ménage », le besoin est davantage tourné vers l'information, voire la promotion de l'enquête auprès du répondant. Et ainsi c'est un portail dit de « *Promotion* » qui est proposé, affichant des informations relatives à une enquête ménage en particulier, ses résultats précédents, son cadre juridique et tout autre élément susceptible de favoriser une réponse par le répondant ménage.

11. Les services de tirage d'échantillon offerts par l'Insee à ses enquêtes ménage et à celles des services statistiques ministériels (SSM) garantissent la non-réinterrogation d'une enquête à l'autre d'un même ménage, afin de limiter la charge supportée individuellement par les répondants.

MUTUALISER SANS DOGMATISME (UNE PARTITION COMMUNE, DES RISQUES DE DISSONANCE)

Cette approche « découplée » est séduisante à plusieurs titres. D'abord, elle offre une modularité intéressante, chaque service pouvant être, au choix, ré-implémenté ou mutualisé. Elle permet également une mise en œuvre incrémentale où la cible métier est atteinte par paliers : plus d'application monolithique « à tout faire », couplée à ses inévitables projets de refonte, mais des ensembles de produits, qui évoluent à des rythmes différents.

C'est ce qui a permis par exemple d'offrir un portail mutualisé à l'enquête Emploi en continu, tout en conservant ses applications de collecte historiques[12] : ainsi, les concepteurs de l'enquête pouvaient mener les travaux méthodologiques nécessaires à la sécurisation du passage au multimode, et ne viser une « bascule » complète qu'avec de nouveaux outils matures. Ou encore, c'est ce qui permet d'assurer la collecte du volet *web* d'enquêtes multimodes, telles que Vie quotidienne et santé (*VQS*) ou Technologies de l'information et des communications (*TIC*), avec du suivi et de l'assistance autour de cette collecte *web*, tout en conservant là aussi les outils historiques pour le volet enquêteur de la collecte.

La cible conceptuelle parfaite ne saurait être complètement atteinte, les habituelles contraintes opérationnelles et budgétaires étant des réalités bien concrètes. Ainsi plus qu'une cible idéale, c'est un cadre, une cohérence globale que les différentes réalisations doivent respecter, des bonnes pratiques qui doivent régir tant la construction des outils que les pratiques métier.

Par exemple, pour les courriers et les différents supports de communication, le cadre initié avec les enquêtes auprès des entreprises a été étendu aux enquêtes auprès des ménages. Ou encore, les services d'assistance offerts aux répondants et leurs processus afférents s'assurent de respecter les mêmes critères de sécurité et de confidentialité des informations personnelles. Ce sont les mêmes outils de conception de questionnaires (*Pogues*, *Eno*) qui sont utilisés, lesquels intègrent déjà les bonnes pratiques méthodologiques.

> *« La mutualisation ou la « généricisation » à outrance ne doit pas devenir un dogme et les réutilisations doivent être étudiées au cas par cas. »*

Mais le pragmatisme implique aussi d'éviter de compliquer inutilement un système déjà inévitablement complexe. La mutualisation ou la « généricisation » à outrance ne doit pas devenir un dogme et les réutilisations doivent être étudiées au cas par cas.

L'existence des deux types de portail des enquêtes de la statistique publique française en est une bonne illustration. Il pouvait paraître séduisant d'imaginer un « méta-portail » proposant des contextes différents selon le type de répondant. Pourtant, une fois correctement découpées en briques fonctionnelles, seules les fonctions d'authentification et d'assistance présentaient des opportunités réelles de mutualisation. Les deux portails couvrent des besoins sensiblement différents et sont des applications simples qui gagnent à le rester, afin de faciliter leur maintenance et leur évolution. Aucun gain structurel significatif n'est à attendre (à ce jour) d'une mutualisation. Là encore, l'approche par les métadonnées offre un angle d'analyse intéressant, des natures de métadonnées différentes, ici de l'information générale sur une enquête contre des liens entreprise-contact et des multiples calendriers d'enquête, suggèrent souvent des besoins « éloignés ».

12. Voir (Guillaumat-Tailliet et Tavan, 2021).

POLY-MODE ET MULTI-MODE, MÊME MUSIQUE ?

Les premiers travaux de conceptualisation pour un passage à l'échelle ont permis de consolider le processus de collecte et de le ré-instancier dans plusieurs contextes : des services du domaine « entreprise » mobilisés dans le contexte ménage, un processus de collecte construit autour du concept d'unité enquêtée, indépendante du mode. On est en mesure de contextualiser les différents concepts, en fonction du mode de collecte.

Mais peut-on simplement mobiliser ce processus pour un mode *web* et pour un mode enquêteur et parler de collecte multimode ? A-t-on résolu pour autant la question du multimode avec cette vision « poly-mode » ?

On pourrait être tenté de répondre oui, mais le multimode ne peut être réduit uniquement à plusieurs processus indépendants. Car il est nécessaire que ces processus interagissent (par exemple, les réponses issues d'un mode doivent pouvoir être reversées dans un autre) et parfois même cohabitent (cas de collecte concurrentielle[13] entre enquêteur et *web*). Se contenter d'une vision poly-mode n'est pas suffisante et il convient de prendre en compte des complexités supplémentaires inhérentes au multimode.

Et complexités supplémentaires entraînent également concepts supplémentaires :

- un questionnement multimode (remobilisation de réponses d'un mode à l'autre, consolidation des réponses en aval, etc.) ;
- un processus multimode (fonctions supplémentaires de suivi multimode, contrôle de la qualité dans un contexte multimode) ;
- des rôles pour les différents acteurs d'une collecte multimode (mobilisation des enquêteurs pour la reprise de questionnaires *web* incomplets, mobilisation de gestionnaires pour la reprise de questionnaires *web* « en erreur », etc).

Il s'agit là encore d'une analyse préalable qui constitue une étape indispensable à toute mise en œuvre d'une collecte multimode. Dans ce cheminement vers un « vrai » multimode, il faut repartir de la question centrale des enquêtes : être en capacité d'exploiter les données collectées.

À CHAQUE TYPE DE DONNÉES SON OUTIL DE CAPTATION (LA FIN DE L'HOMME-ORCHESTRE)

Les données collectées sont l'enjeu principal du processus… de collecte. Celui-ci vise à capter ces données et à permettre leur exploitation. Il s'agit avant tout des données collectées auprès d'un répondant par le biais de questions, mais il existe de fait des données d'autres types :

- des données permettant le suivi et le **pilotage** de la collecte (actions enquêteurs/gestionnaires, plis papier retournés pour adresse invalide, relances envoyées) ;
- des données sur le déroulement de l'enquête : les **paradonnées**, informations techniques captées sur le comportement du répondant (les différents clics, le matériel utilisé, la durée de session, etc.).

13. Les différents types de multi-mode sont déclinés dans l'article déjà cité de François Beck, Laura Castell, Stéphane Legleye et Amandine Schreiber dans ce même numéro.

Les filières « historiques » des enquêtes à l'Insee n'ont pas attendu les réflexions sur le multimode pour mobiliser et collecter ces différents types de données : l'approche strictement mono-mode conduit souvent à produire un questionnaire, un instrument de collecte, avec le *data model*[14] qui sert « d'application à tout faire » (collecter des réponses statistiques, produire des indicateurs de gestions et de suivi, mesurer les temps techniques, qualifier les phases de contact ou la non-réponse, etc.).

Pourtant, les données manipulées ne sont pas destinées aux mêmes acteurs, ni aux mêmes usages. Et une approche découplée préconise qu'elles ne soient pas captées par les mêmes outils. Ainsi, au sein des nouveaux outils pour la mise en œuvre de la collecte multimode :

- les données collectées relèvent de l'instrument de collecte (le questionnaire) ;
- les données de suivi/gestion sont de la responsabilité des outils d'organisation de la collecte (poste de gestion ou poste de collecte enquêteur) ;
- quant aux paradonnées, elles reposent sur des services techniques dédiés et spécialisés.

Dès lors, on est en mesure de proposer une visualisation des données collectées dans un questionnaire ou un autre (reprise de données *web* par un enquêteur par exemple), de permettre de visualiser simplement ou bien de modifier un questionnaire dans un contexte de contrôle qualité sans interférer sur les données de suivi de l'activité enquêteur[15], ou de proposer aux concepteurs d'enquêtes des services de type « entrepôt de paradonnées » afin de supporter des approches expérimentales et des analyses méthodologiques[16].

LE QUESTIONNAIRE «OMNIMODE» (PLUSIEURS COUPLETS, UN MÊME ARRANGEMENT)

Une fois les différents types de données récoltés, il devient possible de traiter l'une des complexités majeures des enquêtes en multimode : leur exploitabilité statistique. Les « effets de mode » – où les réponses à ce que l'on croit être une même question varient en fonction du contexte de questionnement – posent en effet de difficiles questions d'exploitation statistique. Afin d'en limiter l'effet, un certain nombre de principes sont mis en œuvre au sein des protocoles de collecte multimodes[17], mais aussi dans la conception du ou des questionnaires de l'enquête.

> *« Les réflexions sur la portée du questionnaire multimode ne se limitent pas au seul questionnement. »*

Tout d'abord, le questionnement se doit d'être multimode. C'est-à-dire vu comme une consolidation de différents processus de collecte, déclinés opérationnellement selon les modes d'interaction mis en œuvre : les données collectées doivent rester sensiblement les mêmes, les adaptations limitées aux considérations opérationnelles ou « ergonomiques ».

14. En l'occurrence, le *data model* Blaise.
15. Cela évite par exemple d'être obligé de se faire passer pour un enquêteur quand on est gestionnaire, ou de développer une application *ad hoc*.
16. *Généric* est une application de reprise par des gestionnaires, s'appuyant sur des métadonnées actives et la génération automatique de questionnaires. La mise en œuvre d'un entrepôt de paradonnées et du processus de reprise de questionnaires issus du *web* est par ailleurs en cours.
17. Voir l'article de François Beck, Laura Castell, Stéphane Legleye et Amandine Schreiber dans ce même numéro.

Les réflexions sur la portée du questionnaire multimode ne se limitent pas au seul questionnement, ce sont les variables collectées en elles-mêmes et les processus « post-collecte » qui se doivent d'être consolidés.

Pour garantir leur exploitabilité, les données collectées par des outils différents devront mobiliser des chaînes techniques différentes avant d'être enfin mises à disposition dans des bases consolidées. Il y a là aussi une nécessité d'avoir une vision consolidée du processus pour garantir une cohérence : mobiliser les mêmes contrôles post-collecte ou encore les mêmes outils de contrôle de la qualité. Une telle consolidation impose une importante rationalisation du système d'information pour faire « coopérer » les différents modes de collecte.

Les outils doivent permettre la conception d'un questionnaire **omnimode**, en ce sens où il est conçu pour garantir des questionnements les plus similaires possibles d'un mode à l'autre, des valeurs de variables collectées les plus comparables et une exploitation rationalisée.

LE CHAMP DU QUESTIONNAIRE OMNIMODE (SUR TROIS OCTAVES)

Derrière le terme omnimode se cachent en fait trois dimensions :

- le questionnaire est « unique » et commun à tous les modes ;
- le questionnaire est adapté à tous les modes ;
- les réponses issues des différents modes doivent être réconciliées en une même variable statistique.

Les deux premières dimensions sont largement prises en compte dans les outils de conception de questionnaires déjà mis en œuvre à l'Insee : *Pogues* et *Eno* permettent d'adapter automatiquement les mêmes objets de questionnaire (questions, champ de réponse, etc.) aux différents modes de collecte et ils limitent les objets spécifiques à un mode de collecte donné.

La déclinaison d'objets du questionnaire aux différents modes de collecte a déjà été évoquée, quant aux éléments réservés à un seul mode, c'est une pratique que les outils permettent mais limitent. La contextualisation d'une consigne (destinée à un enquêteur ou destinée à un répondant *web* par exemple) est assez courante et usuelle, mais filtrer des questions ou blocs de question selon le mode doit être encadré pour préserver le principe de questionnaire omnimode avec les mêmes variables collectées. La bonne pratique consiste à limiter ces cas aux questions dont les réponses peuvent être imputées automatiquement, garantissant une certaine indépendance des traitements post-collecte (chaque réponse collectée passe par les mêmes chaînes de traitement quel que soit le mode).

L'exemple théorique de la question « Avez-vous, personnellement, la possibilité d'accéder à un internet ? »[18] en est une bonne illustration. Dans un contexte *web*, elle paraît superflue et la réponse pourrait être imputée automatiquement à « Oui ». On peut également citer les questions relatives à la description d'un logement (distinction entre appartement ou maison individuelle par exemple) qui peuvent relever des phases de repérage préalables par l'enquêteur et alléger d'autant le questionnement statistique en face-à-face (tout en restant pertinentes dans les autres modes de collecte).

18. Au-delà des considérations de formulation, elle est vue comme un cas d'école exemple de question « absurde » dans le contexte *web*.

Pour ce qui est de la troisième dimension du questionnaire omnimode, des variables statistiques consolidées, elle constitue un champ d'évolution à considérer. Il semblerait efficace de pouvoir spécifier tant la variable statistique, que son mode de collecte (la question, son libellé, etc.), que les contraintes de format (les différents contrôles dans les questionnaires) et aller jusqu'aux règles de consolidation ou d'imputation automatique dont elle doit faire l'objet post-collecte.

Ainsi, en reprenant l'exemple didactique précédent, en même temps que le libellé de la question « Avez-vous personnellement accès à internet ? », une règle d'imputation automatique à « Oui » pour les répondants issus du mode *web* pourrait être spécifiée. Ou encore la spécification d'une variable à valeur unique prise dans une liste de modalité (une « liste déroulante » sur internet) pourrait être enrichie des règles de redressement et de consolidation multimode : on pense à l'exemple classique de la réponse exclusive « oui/non » pour laquelle un répondant facétieux peut cocher les deux réponses sur un questionnaire papier ; à l'occasion de la conception de la question, une règle de redressement automatique pourrait également être spécifiée (mettre à blanc ou privilégier le « Oui » en cas de réponses multiples non souhaitées).

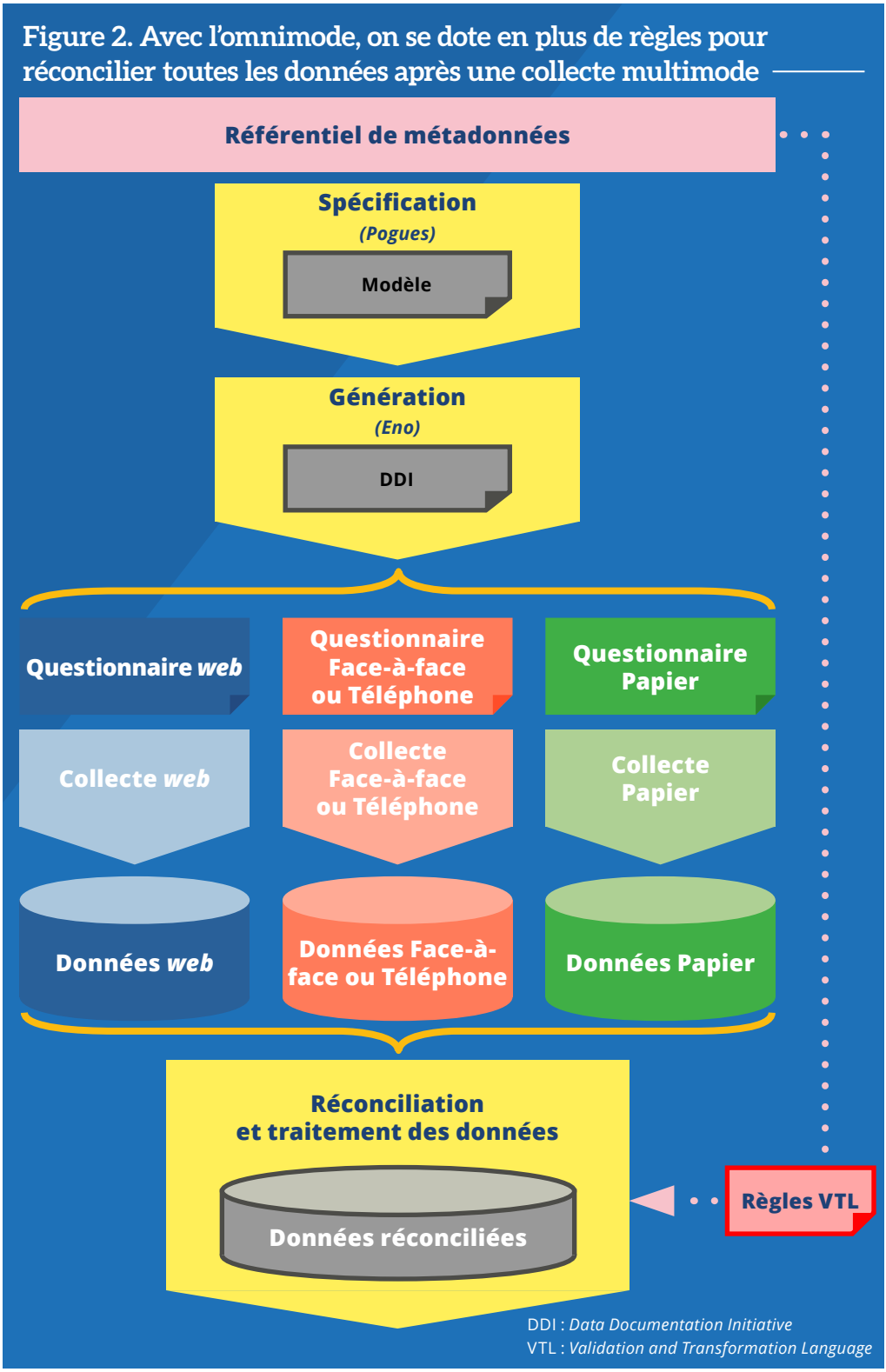
Ce type de traitement est actuellement réalisé *a posteriori* dans le traitement en aval de la collecte, et de manière spécifique à chaque enquête. Une règle d'imputation de ce type nécessite un certain formalisme d'écriture, et viendrait enrichir les éléments de spécification du questionnaire pour décrire ces traitements tout au long du processus. L'Insee a engagé des travaux autour de VTL, méta-langage de spécification pour les règles algorithmiques, promu par Eurostat : on spécifie dans le même langage, les filtres, les contrôles en cours de collecte ou post-collecte, les variables calculées au sein du questionnaire, ou les variables statistiques à livrer[19]. Et on « active » au passage de nouvelles métadonnées.

> *« Plus qu'un simple questionnement, il s'agit de concevoir la captation de variables statistiques, un « arbre de variables » dont on souhaite disposer. »*

À partir d'une conceptualisation omnimode de ce que doit être le questionnement statistique, on limite les effets de mode, on améliore l'exploitabilité statistique des résultats et on rationalise les outils mobilisés dans la consolidation des variables collectées. Ainsi, plus qu'un simple questionnement, il s'agit de concevoir la captation de variables statistiques, un « arbre de variables » dont on souhaite disposer, demandant l'instanciation de processus de collecte multiples, et l'ensemble des réponses collectées devant se « réconcilier » dans une base consolidée ***(figure 2)***.

19. Cette spécification consolidée peut se décliner ensuite automatiquement dans les différents outils et traitements bien que dans des technologies informatiques différentes (Java, JavaScript, etc.). Voir (SDMX, 2020 ; Banca d'Italia, European Central Bank et Insee, 2021).

Figure 2. Avec l'omnimode, on se dote en plus de règles pour réconcilier toutes les données après une collecte multimode



LES ENJEUX DU PROCESSUS MULTIMODE (DES ACTEURS QUI DOIVENT AGIR DE CONCERT)

Une collecte multimode, s'appuyant sur un questionnaire omnimode pose encore un certain nombre de problématiques opérationnelles et organisationnelles.

> *« Découpler les enjeux méthodologiques et statistiques d'un concepteur et ceux plus opérationnels d'un maître d'œuvre du processus de collecte. »*

Quelle charge représente pour un enquêteur un questionnaire collecté par internet ? Un questionnaire repris d'internet ? Comment correctement coordonner l'action d'un enquêteur quand il y a des relances papier de masse ? Comment répercuter des événements complexes tels que les éclatements de ménage, de logement ou de budget sur les différents processus ? Comment organiser correctement le travail des enquêteurs quand de tels évènements proviennent de réponses collectées par internet ?

La qualité statistique obtenue dépend tout autant de la conception rigoureuse du questionnement que de la mise en œuvre du processus de collecte où chaque acteur a son rôle à jouer. Avec l'arrivée du multimode, c'est tout un processus et les rôles et responsabilités afférents des différents acteurs qui est à penser. Et là encore, découpler les enjeux méthodologiques et statistiques d'un concepteur de questionnement (questions correctement comprises, variables d'intérêt identifiées, protocole maîtrisé) et ceux plus opérationnels d'un maître d'œuvre du processus de collecte (suivi d'activité des enquêteurs, gestion du réseau enquêteur, priorisation fine d'opérations) relève d'une étape préalable. Leurs objectifs, sans être divergents, gagnent à être respectivement identifiés et consolidés.

UN NOUVEAU CONCEPT ? L'UNITÉ D'INTÉRÊT D'ENQUÊTE (UNE NOUVELLE PORTÉE)

La maîtrise d'ouvrage de l'enquête, qui conçoit une enquête, son protocole, son questionnaire, son calendrier, etc. s'intéresse avant tout à l'unité statistique « initiale » ou « finale ».

C'est le concept d'**unité d'intérêt d'enquête**[20] qui est ici introduit : il correspond à l'unité de compte statistique issue de l'échantillonnage initial. Il garantit que les données liées à cette unité seront correctement collectées (pas de doublon, pas de « trous de collecte ») et correctement livrées (une réconciliation technique sous forme de donnée statistique avant la correction d'effet de modes éventuels).

Le maître d'œuvre statistique doit pour sa part organiser le travail des différents acteurs, et mettre en œuvre les outils nécessaires, afin de permettre le bon déroulement des différents processus de collecte, parfois concomitants, sur plusieurs enquêtes. Il a besoin, lui, d'une unité de collecte, l'**unité enquêtée**, qui va lui garantir que les événements, les coûts et charges induits, liés au déroulement du processus de collecte d'une unité seront correctement captés pendant les opérations.

Sans cette dualité de concept, les informations orientées « contrôle de gestion » sont mélangées avec celles plus orientées « statistique » et les objectifs de chacune de ces deux approches sont dévoyés.

20. Qu'elle fasse l'objet d'une collecte multimode concurrentielle ou séquentielle, elle reste une unité de compte unique, à laquelle peuvent être rattachées plusieurs unités enquêtées, dans des modes différents, à des moments différents. Elle peut formellement être rapprochée du concept d'unité d'intérêt statistique du GSIM (*General Statistical Information Model*).

Un exemple caractéristique est celui de la filière historique Insee où étaient produits des « codes résultats ». Ces agrégats étaient calculés en cours de collecte afin d'offrir tant des éléments pour le calcul de la performance des enquêteurs que ceux nécessaires au redressement, imputation et autres traitements statistiques dit « aval ». La liste des « codes résultats » existants était devenue complexe, difficile à maintenir, difficile à consolider entre les enquêtes et engendrait parfois des compromis imposés quant au besoin statistique ou opérationnel couvert. L'ambiguïté sur le terme de « hors-champ » en est une bonne illustration :

- il peut être entendu au sens opérationnel du terme – un logement détruit par exemple – et n'être comptabilisé ni dans le calcul de performance des enquêteurs, ni dans les résultats statistiques ;
- ou entendu au sens statistique exclusif – un ménage sans fonctionnaire pour une enquête sur la fonction publique d'état comme autre exemple – et pourtant devant être comptabilisé comme une enquête réussie pour l'enquêteur.

Pour répondre aux deux besoins, il convient de découpler les concepts, et surtout de déporter le calcul des agrégats dans les processus qui en sont consommateurs : l'un pour charger un infocentre utile au contrôle de gestion sur l'activité des enquêteurs, l'autre pour alimenter les chaînes de traitements avec la qualification de la réponse, ou plus souvent de la non-réponse, statistique. À l'instar des deux portails adaptés aux deux types de répondants aux enquêtes, ce découplage permet de limiter la complexité induite et de garantir que chacun voit son besoin couvert, chacun étant responsable de ses propres agrégats.

Aujourd'hui, les outils que l'Insee met à disposition des enquêteurs implémentent ces principes. Par exemple, les fonctions d'organisation et de suivi de la collecte sont dans des « briques » applicatives séparées des fonctions liées à la conduite d'un questionnement statistique.

SAVOIR ET POUVOIR S'ENTENDRE ENTRE MODES (POUR ÉVITER LA CACOPHONIE)

La dernière étape de conceptualisation, non la moindre, pour être en mesure de mettre en œuvre correctement une collecte multimode concerne la « communication multimode ».

Un processus multimode permet de mettre en œuvre une collecte quel que soit le mode. Une consolidation au niveau du questionnement statistique garantit une exploitabilité statistique. Une consolidation au niveau de l'unité de compte statistique garantit un pilotage cohérent et efficace des collectes successives d'une enquête. Mais une dimension supplémentaire n'est pas encore prise en compte : le besoin des processus de communiquer « en direct » les uns avec les autres.

Cette communication est indispensable pour permettre une collecte multimode concurrentielle ou à séquences rapprochées (par exemple, les processus « collecte *web* » et « collecte enquêteur » devant s'avertir l'un l'autre), pour permettre un suivi consolidé des différentes collectes (le suivi des différentes unités enquêtées, quel que soit leur mode de collecte, doit être cohérent et consolidé dans une même interface), ou encore la mise en œuvre de collectes hybrides (comme celles à base de carnet de collecte, où le démarrage de la collecte *web* du carnet doit être synchronisé avec le passage d'un enquêteur).

Dans une collecte multimode « totale », on peut facilement imaginer trois à quatre processus de collecte simultanés devant communiquer les uns avec les autres. Ces différents processus en forment-ils un seul complexe, ou plusieurs se « parlant » ? Comment s'assurer que cette communication ne va pas devenir effectivement cacophonique ? En clair, comment doivent être conçues ces collectes parallèles ?

LES DIFFÉRENTS MODES DE COLLECTE À L'UNISSON

À l'Insee cette conception a été assez « organique ». Partant d'une conceptualisation initiale poly-mode, où un processus de collecte peut être mis en œuvre plusieurs fois dans des modes différents, de premières opérations d'enquêtes ont pu être conduites. Ainsi quand la délicate question de faire communiquer ces processus s'est posée – par exemple quand une collecte internet doit « prévenir » un enquêteur qu'une réponse a été reçue, ou encore quand un enquêteur doit « prévenir » un carnet *web* de collecte qu'un nouveau répondant va arriver – l'enjeu était de préserver la stabilité des systèmes existants, comme de pérenniser les opérations de collecte déjà mises en œuvre.

Dans une approche « mono-processus », un processus complexe avec plusieurs instruments de collecte, plusieurs règles de gestion et de synchronisation, permet de gérer toutes les collectes, de la plus simple ne mobilisant qu'un mode de collecte, jusqu'aux multimodes. Plutôt que d'opter pour cette approche, le choix s'est porté sur une approche « poly-processus », où plusieurs processus cohabitent indépendamment les uns des autres et communiquent avec un système central ***(figure 3)***. Ce système est alors le chef d'orchestre de la collecte multimode, seul porteur de la complexité des règles de gestion liées aux protocoles (quels modes faut-il mobiliser, quelles actions déclencher, etc.).

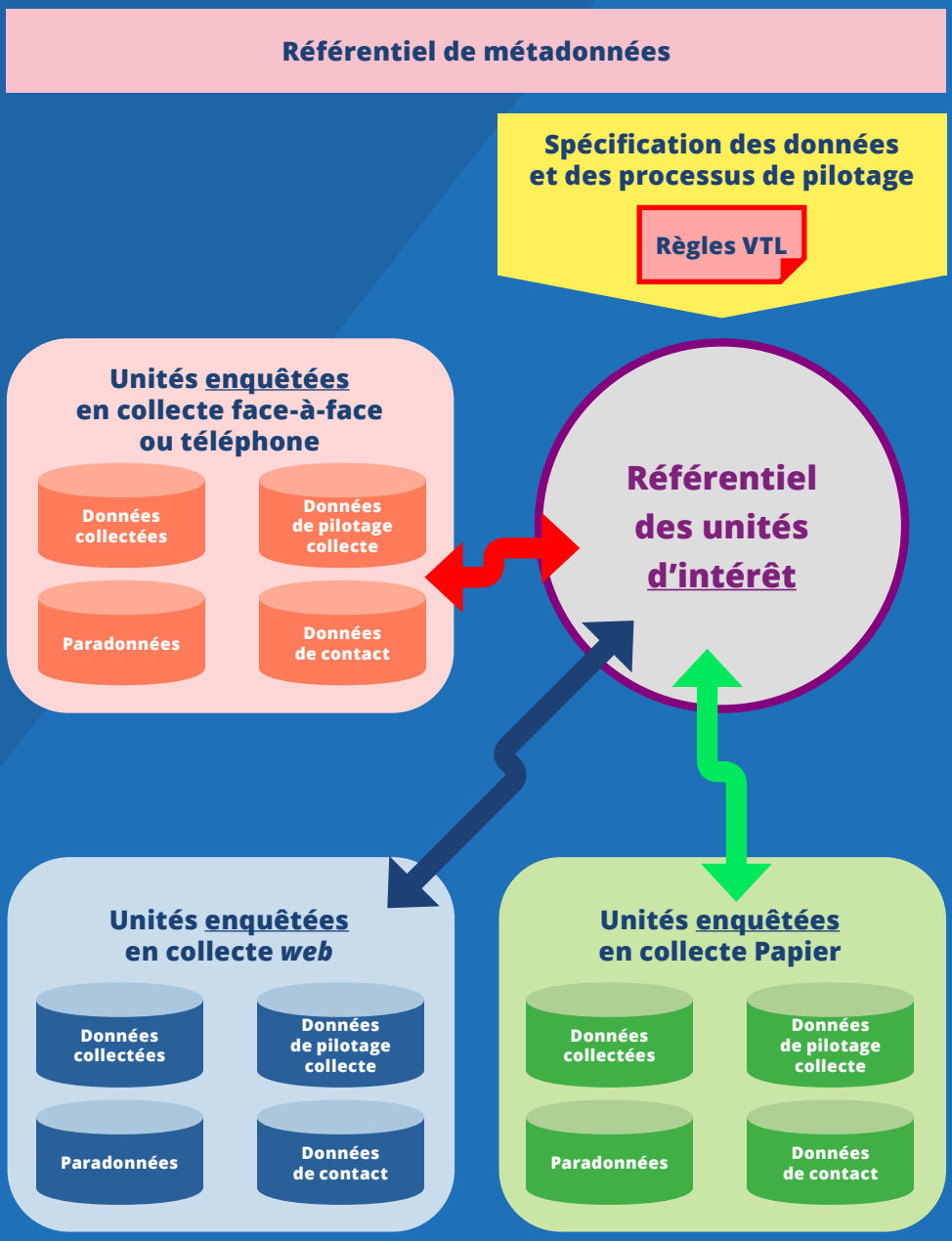
Ainsi, la montée en charge et la complexité croissante des protocoles considérés ont des impacts limités sur la complexité de chaque processus. D'une part, seuls les protocoles de collecte multimode les plus complexes requièrent une réelle communication inter-mode (collecte concurrentielle pour les chiffres de l'emploi, ou collecte s'appuyant sur des carnets d'activité ou de budget notamment). La grande majorité des collectes, mêmes multimodes, sont séquentielles et un modèle « poly-processus » convient parfaitement. D'autre part, c'est le système central, en charge de la communication entre chacun des processus, qui est le plus impacté par la complexité à mettre en œuvre. Les évolutions à porter dans les systèmes existants pour supporter les protocoles les plus complexes sont ainsi limitées (schématiquement des échanges de « messages » avec un système central).

PREMIÈRES OPÉRATIONS MONO-MODES (AVEC DES TONALITÉS MULTIMODES)

Les premières étapes du passage à l'échelle, conforme avec le cadre conceptuel défini, visent à sécuriser des systèmes de collecte indépendants. Une première collecte par internet dans un contexte ménage, une première collecte par téléphone, une première collecte par face-à-face.

Si ces premières opérations ne mobilisent pas de « communication » multimode, à proprement parler, la dimension multimode est déjà présente dans le processus et le travail des différents acteurs. Par exemple, la première collecte par téléphone se déroule dans une opération multimode séquentielle, le pilote 2022 de la nouvelle enquête Logement.

Figure 3. Le pilotage d'une collecte multimode génère un fort besoin de synchronisation



VTL : *Validation and Transformation Language.*

L'unité enquêtée est l'unité de compte d'une opération de collecte.

L'unité d'intérêt est l'unité de compte statistique.

À une unité d'intérêt peut correspondre plusieurs unités enquêtées (par exemple si un questionnement web s'enchaîne avec une collecte par téléphone pour un même ménage ou une même entreprise.).

Dans des collectes séquentielles, où le premier mode est le *web*, des stratégies d'optimisation des relances et de bascule entre les modes sont mises en œuvre. Elles nécessitent que les acteurs maîtrisent les différents processus et puissent manipuler les différents types de données (données collectées, paradonnées, données de suivi) même si les outils ne communiquent pas formellement directement.

Dans ces mêmes collectes, mobiliser des réponses antérieures issues d'internet dans une collecte par enquêteurs, nécessite que soit précisé le cadre méthodologique. Attend-on de l'enquêteur une reprise exhaustive du questionnaire, les réponses internet servant de « fil rouge » à un entretien classique ? Ou au contraire une réponse efficace, les réponses internet servant de « pis-aller » pour permettre un entretien raccourci ? La formation des enquêteurs, la prise en compte de ce nouveau métier dans leurs tâches fait partie intégrante des impacts du multimode, et ce dès les premières opérations réputées « mono-modes ».

Plus généralement, décliner et formaliser les protocoles multimodes en consignes, tâches ou rôles des acteurs intervenants dans la collecte (de l'enquêteur, à la maîtrise d'ouvrage d'enquête ou la maîtrise d'œuvre statistique) est l'un des enjeux de ces premières opérations.

GAGNER EN VIRTUOSITÉ EN ABORDANT UN MULTIMODE DE PLUS EN PLUS COMPLEXE

Avec la mise en œuvre de collectes multimodes plus complexes, ce ne sont plus seulement les applications, les outils ou les systèmes techniques qui doivent évoluer, mais aussi les métiers.

La méthode d'analyse préalable des processus cible est encore à l'œuvre, avec son lot de questions :

- Qu'est-ce que le suivi multimode ? Sur quels indicateurs s'appuyer pour suivre une collecte consolidée ?
- Qu'est-ce que le contrôle qualité des questionnaires internet dans la sphère ménage ? Faut-il se calquer sur un processus type « entreprise », s'appuyant sur des gestionnaires ? Ou faire intervenir l'expertise de l'enquêteur ?
- Combien de temps prend effectivement la reprise d'un questionnaire internet par un enquêteur ? Plus longtemps, car il y a un temps de préparation préalable et une phase d'accompagnement du répondant un peu plus longue ? Ou moins longtemps, car l'entretien est raccourci ?

Autant de questions complexes qui ne sauraient accepter des réponses *a priori*. L'expérimentation sera donc encore de mise, pour pouvoir tester et constater dans les faits. Cela nécessitera que les processus conçus et que les outils proposés soient suffisamment souples : des services standardisés faciles à mettre en œuvre, des acteurs formés et sensibilisés au caractère expérimental. Car dans un système de collecte multimode, la complexité impose de la souplesse, pour que s'approfondisse la connaissance de ce nouveau métier et que les interprètes des prochaines collectes s'approprient progressivement les techniques de demain.

BIBLIOGRAPHIE

BANCA D'ITALIA, EUROPEAN CENTRAL BANK et INSEE, 2021. *VTL Community. Towards a community of VTL developers.* [en ligne]. GitHub. [Consulté le 6 décembre 2021]. Disponible à l'adresse : *https://github.com/VTL-Community/VTL-Community*.

BONNANS, Dominique, 2019. RMéS, le référentiel de métadonnées statistiques de l'Insee. In : *Courrier des statistiques*. [en ligne]. 27 juin 2019. Insee. N° N2, pp. 46-57. [Consulté le 6 décembre 2021]. Disponible à l'adresse : *https://www.insee.fr/fr/statistiques/fichier/4168396/courstat-2-6.pdf*.

COTTON, Franck et DUBOIS, Thomas, 2019. Pogues, un outil de conception de questionnaires. In : *Courrier des statistiques*. [en ligne]. 19 décembre 2019. N° N3, pp. 17-28. [Consulté le 6 décembre 2021]. Disponible à l'adresse : *https://www.insee.fr/fr/statistiques/fichier/4254216/courstat-3-3.pdf*.

COTTON, Franck, DUBOIS, Thomas, SIGAUD, Éric et WERQUIN, Benoît, 2021. *A fully metadata-driven platform for the conception of survey questionnaires and the management of multimode data collection.* [en ligne]. 27-30 septembre 2021. Unece, Conférence des statisticiens européens, Expert Meeting on Statistical Data Collection. [Consulté le 6 décembre 2021]. Disponible à l'adresse : *https://unece.org/sites/default/files/2021-09/DC2021_S1_France_Werquin%20Cotton_AD.pdf*.

ERIKSON, Johan, 2020. Le modèle de processus statistique en Suède – Mise en œuvre, expériences et enseignements. In : *Courrier des statistiques.* [en ligne]. 29 juin 2020. Insee. N° N4, pp. 122-141. [Consulté le 6 décembre 2021]. Disponible à l'adresse : *https://www.insee.fr/fr/statistiques/fichier/4497085/courstat-4-8.pdf*.

GUILLAUMAT-TAILLIET, François et TAVAN, Chloé, 2021. Une nouvelle enquête Emploi en 2021, entre impératif européen et volonté de modernisation. In : *Courrier des statistiques*. [en ligne]. 8 juillet 2021. Insee. N° N6, pp. 7-27. [Consulté le 6 décembre 2021]. Disponible à l'adresse : *https://www.insee.fr/fr/statistiques/fichier/5398681/courstat-6-art-1.pdf*.

HAAG, Olivier et HUSSEINI-SKALITZ, Anne, 2019. Collecte par internet des enquêtes auprès des entreprises : l'Insee entre dans l'ère Coltrane. In : *Courrier des statistiques*. [en ligne]. 19 décembre 2019. N° N3, pp. 45-60. [Consulté le 6 décembre 2021]. Disponible à l'adresse : *https://www.insee.fr/fr/statistiques/fichier/4254223/courstat-3-5.pdf*.

KOUMARIANOS, Heïdi et SIGAUD, Éric, 2019. Eno, un générateur d'instruments de collecte. In : *Courrier des statistiques*. [en ligne]. 19 décembre 2019. N° N3, pp. 29-44. [Consulté le 6 décembre 2021]. Disponible à l'adresse : *https://www.insee.fr/fr/statistiques/fichier/4254218/courstat-3-4.pdf*.

SDMX, 2020. *Validation and Transformation Language (VTL).* [en ligne]. Mise à jour du 4 août 2020. The official site for the SDMX community. A global initiative to improve Statistical Data and Metadata eXchange. [Consulté le 6 décembre 2021]. Disponible à l'adresse : *https://sdmx.org/?page_id=5096*.

UNECE, 2019. *Generic Statistical Business Process Model GSBPM.* [en ligne]. Janvier 2019. Version 5.1. [Consulté le 6 décembre 2021]. Disponible à l'adresse : *https://statswiki.unece.org/display/GSBPM/GSBPM+v5.1*.