



Marco D'Angelo, 05.03.2020

Directives sur l'appariement

Version:	Date:	Motif:
0.1	22.05.2015	Arbeitsgruppensitzung vom 22.05.2015 (Titel „Aussprachepapier“)
0.2	08.06.2015	Arbeitsgruppensitzung vom 10.06.2015 (Titel „Aussprachepapier“)
0.3	17.06.2015	Arbeitsgruppensitzung vom 23.06.2015 (Titel „Aussprachepapier“)
0.9	24.06.2015	GL-Sitzung vom 01.07.2015 (Titel „Aussprachepapier“)
0.91	22.09.2015	Erster Entwurf für Arbeitsgruppensitzung vom 24.09.2015
0.93	24.09.2015	Entwurf für Arbeitsgruppensitzung vom 24.09.2015
0.94	29.09.2015	GL-Sitzung vom 07.10.2015
0.95	29.09.2015	GL-Sitzung vom 07.10.2015 (Version mit Änderungen)
0.96	27.10.2015	Version für Arbeitsgruppensitzung vom 19.11.2015
0.97	19.11.2015	Version für das BFS-Kolloquium vom 25.11.2015
1.1	08.11.2016	Version für Arbeitsgruppensitzung vom 20.12.2016 (Freigabe)
1.2	05.03.2020	Version avec adaptations aux nouvelles conditions-cadres

Destinataires: GL, groupe de travail Appariements

Glossaire

Abréviation	Signification
CD	Chef/cheffe de division de l'OFS
OFS	Office fédéral de la statistique
OFIT	Office fédéral de l'informatique et de la télécommunication
REE	Registre des entreprises et des établissements
LSF	Loi sur la statistique fédérale
COP	Code of Practice
DFI	Département fédéral de l'intérieur
SPROD	Section de production de l'OFS
DC	Données consolidées
PFPDT	Préposé fédéral à la protection des données et à la transparence
ETL	Extract, Transform, Load (Process)
GL	Conseil de direction de l'OFS
RegBL	Registre des bâtiments et des logements
ISDS	Sécurité de l'information et protection des données
IT	Section de l'informatique de l'OFS
METH	Section des méthodes statistiques de l'OFS
PPA	Programme pluriannuel de l'OFS
OSS	On Site Support (de l'OFIT)
RD	Service juridique de l'OFS
LHR	Loi fédérale sur l'harmonisation des registres des habitants et d'autres registres officiels de personnes
ESPA	Enquête suisse sur la population active
CS	Chef/cheffe de section de l'OFS
RS	Relevé structurel
SESAM	Enquête de synthèse sur la protection sociale et le marché du travail
SHIVALV	Aide sociale, assurance-invalidité et assurance-chômage
SHS	Statistique de l'aide sociale
SILC	Enquête sur les revenus et les conditions de vie
SIS	Système d'information statistique
SMS	Statistical Metadata System
SUS	Statistique des condamnations pénales
STATENT	Statistique structurelle des entreprises
STATPOP	Statistique de la population et des ménages
OLPD	Ordonnance relative à la loi fédérale sur la protection des données
VS	Service spécialisé en appariements de l'OFS

Table des matières

1	Résumé	5
2	Contexte	5
3	Buts des directives	6
3.1	D'une approche axée sur les enquêtes à une approche axée sur les résultats	6
3.2	Protection et sécurité des données	6
4	Lien avec le programme pluriannuel de la statistique et buts de l'office, des divisions et des sections	6
5	Cadre légal	7
6	Bases de l'appariement	8
6.1	Définitions	8
6.1.1	Qu'est-ce qu'un appariement?	8
6.1.2	Catégories d'appariements	9
6.2	Principes applicables	12
6.3	Prescriptions relatives à l'appariement	13
6.4	Sensibilité des données	17
7	Production statistique, sécurité et appariement des données	18
7.1	Processus de production statistique	18
7.2	Protection des données	19
7.2.1	Espace de confiance à l'OFS	20
7.2.2	Secteurs	20
7.2.3	Pseudonymisation des données consolidées	20
7.3	Principe de la gestion des clés	22
8	Organisation et rôles	22
9	Éléments de l'appariement de données	24
9.1	Inventaires des appariements à l'OFS	24
9.2	Demandes d'appariement	25
9.3	Règlement de traitement	26
10	Processus d'appariement	28
10.1	Remarques préliminaires	28
10.2	Examen de la demande et décision	28
10.2.1	Examen de la demande	28
10.2.2	Décision de la direction	29
10.3	Réalisation de l'appariement	30
10.3.1	Appariement par l'OFS	30
10.3.2	10.3.2 Association de tiers au processus d'appariement	31
10.3.3	Appariement par des services cantonaux ou communaux	32
10.4	Remise de données appariées	32
10.5	Clôture du projet d'appariement et effacement des données appariées	32
11	Annexes	34
11.1	Annexe 1: Définitions (concept SIS)	34
11.2	Annexe 2: Bases légales	36
11.3	Annexe 3: Règlement de traitement (modèle)	39

Directives sur l'appariement

11.4	Annexe 4: Description des processus (voir aussi chap. 10)	40
11.4.1	Traiter une demande d'appariement (vue d'ensemble).....	40
1.	Spécifier la demande	40
2.	Prendre position sur la demande	41
3.	Décider de la demande.....	41
4.	Préparer l'appariement	42
4.1	Générer une clé de pseudonymisation „SIS“.....	42
4.2	Utiliser un pc « appariement »	42

1 Résumé

L'appariement de données vise à produire des informations nouvelles à partir des données existantes, tout en évitant des relevés superflus, en allégeant la charge des milieux interrogés, en créant des synergies et en définissant des sources de données élargies. Il consiste à relier entre elles à des fins statistiques des données individuelles qui proviennent de différentes sources, en créant ainsi un nouveau jeu de données. A cet égard, on distingue quatre types d'appariement: les appariements systématiques destinés à la production statistique, les appariements longitudinaux, les appariements destinés à l'analyse statistique et les appariements spéciaux. Un inventaire des appariements est dressé afin d'en garantir la traçabilité.

Les appariements effectués à l'OFS satisfont à des conditions très strictes de protection et de sécurité des données. Tous les appariements font l'objet d'une demande d'approbation et doivent recevoir l'aval du directeur. Différents secteurs sont définis au sein de l'OFS. Ces secteurs, qui correspondent aux unités de production statistique, visent l'obtention de données consolidées. Seules les personnes autorisées ont accès aux données sectorielles (données consolidées), et seuls les détenteurs des clés de dé-pseudonymisation peuvent appairer des données. L'utilisation de ces clés n'est possible que si l'appariement a été approuvé; elle est enregistrée dans un journal. Lorsque le jeu de données qui doit résulter de l'appariement atteint le degré de protection 3, il est nécessaire d'établir un règlement de traitement.

Deux organes ont été constitués pour encadrer les projets d'appariement: le service spécialisé en appariements, qui coordonne les travaux administratifs et qui gère les clés, et le groupe de travail Appariements.

Des appariements peuvent être demandés aussi bien par des services de l'OFS que par des services externes. C'est en premier lieu l'OFS qui se charge des appariements. Les cantons, les communes et des externes peuvent être autorisés à appairer des données de l'OFS à certaines conditions.

2 Contexte

Les travaux sur l'appariement des données menés depuis 2010¹ ont débouché sur une conception comprenant trois volets:

- Orientation générale
- Bases légales
- Organisation/processus/informatique

Un règlement de traitement a été élaboré en 2014 (règlement de traitement Appariement, version du 13 octobre 2014). Soumis au préposé fédéral à la protection des données et à la transparence (PFPDT), il a reçu de ce dernier un préavis favorable. Les travaux n'étaient cependant pas terminés. Le projet a été présenté au conseil de direction en avril 2015. Le conseil de direction s'est appuyé sur cette présentation pour formuler un cahier des charges. Il s'est penché en juillet 2015 sur une note de discussion basée sur le règlement de traitement de 2014. Ce dernier a été révisé, puis approuvé par le conseil de direction en décembre 2015, puis intégré aux présentes directives.

En 2018, la manière de travailler dans le domaine des appariements de données a été expliquée au PFPDT lors d'une visite à l'OFS. Le PFPDT a pris connaissance de la documentation établie par l'OFS.

Lors de visites à l'OFS, les organes Fedestat et Corstat ont eux aussi été informés que les appariements sont intégrés dans le système statistique public de la Suisse.

¹ Note de discussion 2010, version 1.0 du 25.8.2010

3 Buts des directives

3.1 D'une approche axée sur les enquêtes à une approche axée sur les résultats

La plupart des statistiques étaient auparavant fondées sur les relevés: chaque statistique faisait l'objet d'un relevé ciblé. La production statistique a été adaptée de manière à gagner en efficacité, dans le but d'éviter les doublons, de décharger les personnes interrogées et de mieux exploiter les synergies. Conformément à l'art. 4 LSF, les données statistiques doivent être obtenues en premier lieu sans réaliser de relevés directs, autrement dit en recourant de manière conséquente à des registres et à des données administratives. Cela passe par des systèmes d'informations intégrés, qui sont axés sur l'output et qui se fondent sur des relevés coordonnés de données provenant de différentes sources (registres, données administratives ou enquêtes). Les données sont appariées (combinées) et exploitées pour produire l'output souhaité (statistiques). Un tel système a comme avantage, en plus de favoriser le recyclage des données saisies, de permettre la réalisation de nouvelles statistiques à partir des données recyclées. Grâce à cette souplesse, on peut mieux satisfaire les besoins en informations croissants des milieux économiques et politiques, de la société, de l'administration et de la recherche.

3.2 Protection et sécurité des données

La loi sur la statistique fédérale, la loi sur la protection des données, le code de bonnes pratiques d'Eurostat, la charte de la statistique publique et les directives concernant la sécurité informatique dans l'administration fédérale s'appliquent de manière générale à l'OFS. L'office applique les mesures suivantes pour garantir la protection et la sécurité des données:

- authentification forte (accès au moyen de la Smartcard)
- limitation des autorisations d'accès
- signature par chaque membre du personnel d'une déclaration par laquelle il accepte de se conformer aux règles de protection des données
- anonymisation des données individuelles remises
- conclusion d'un contrat de protection des données avec les utilisateurs externes

D'autres mesures sont nécessaires pour garantir notamment la transparence des appariements et des processus, la protection des données et la pseudonymisation des données individuelles.

Les éléments suivants doivent être réglementés afin de satisfaire aux exigences et de garantir que l'appariement se fasse de manière contrôlée, sûre et transparente:

- définitions (chap. 6.1)
- principes régissant l'appariement (chap. 6.2)
- conditions de l'appariement (chap. 6.3)
- dispositions techniques et gestion des données (chap. 7)
- organisation, rôles et compétences (chap. 8)
- processus d'appariement (chap. 10).

4 Lien avec le programme pluriannuel de la statistique et buts de l'office, des divisions et des sections

L'objectif 3 du programme pluriannuel de la statistique fédérale PPA 2015 – 2019 souligne l'importance de l'appariement. Il dispose que «la statistique fédérale exploite les potentiels qui ont été créés pour produire des informations statistiques nouvelles et adapte la diffusion des informations statistiques aux besoins nouveaux des utilisateurs. L'appariement est depuis lors considéré comme un élément prioritaire dans le PPA.

Directives sur l'appariement

Les objectifs de l'OFS intègrent l'objectif du programme pluriannuel susmentionné. Le groupe de travail Appariement poursuit l'objectif suivant: exploiter le potentiel de l'appariement des données selon les processus et les structures définis et optimiser constamment ce potentiel par le recours aux bonnes pratiques et les échanges avec les divisions.

5 Cadre légal

Les bases légales applicables sont notamment l'art. 14a LSF, l'art. 16, al. 4, LHR, l'art. 21 OLPD, l'art. 13h ss de l'ordonnance sur les relevés statistiques, l'ordonnance du DFI sur l'appariement de données et la présente directive de l'OFS sur l'appariement de données. Ces dispositions définissent le cadre légal suivant:

Raison d'être: Les relevés directs peuvent être fastidieux pour les personnes interrogées et suscitent parfois chez ces dernières ire et incompréhension vis-à-vis de la Confédération (et plus particulièrement de l'OFS); les taux de réponses en souffrent, et la qualité des statistiques s'en trouve affectée à son tour. Pour pallier ce problème, l'art. 4, al. 3, LSF dispose que le nombre des enquêtes doit être limité au strict nécessaire. Pour préserver la qualité et la représentativité de ses statistiques, l'OFS dispose de l'outil de l'appariement.

Compétence: Au niveau fédéral, seul l'OFS est habilité à effectuer des appariements, afin de remplir ses tâches statistiques. Il s'agit d'une habilitation d'ordre général (art. 14a LSF). L'art. 16a, al. 4, LHR autorise par ailleurs l'appariement de données des registres des habitants avec des données du RegBL et du REE.

Les services statistiques cantonaux et communaux peuvent appairer des données de l'OFS pour remplir leurs tâches statistiques à condition qu'ils aient l'autorisation écrite de l'OFS et qu'ils satisfont à ses exigences.

D'autres tiers (mandants) peuvent être associés au processus d'appariement pour réduire les coûts et le travail engendrés. Les détails doivent si nécessaire être clairement définis dans le contrat de protection des données.

Transparence: Une demande écrite et motivée doit être présentée pour chaque appariement. Cette demande doit être soumise aux sections de production concernées, à la section Méthodes statistiques et au service juridique, qui en vérifient la faisabilité sur les plans technique, méthodologique et juridique. L'annexe de l'ordonnance sur les relevés statistiques fédéraux précise les statistiques qui sont réalisées au moyen d'appariements systématiques.

Protection des données: Les appariements de données ne sont autorisés que s'ils sont nécessaires et adaptés aux travaux statistiques et qu'ils présentent la qualité requise. Font exception les cas spéciaux internes auxquels des critères spécifiques sont appliqués et qui servent à tester la qualité des données. Le service d'appariement doit centraliser et conserver de manière sûre les clés nécessaires à la réalisation des appariements, conformément aux principes de gestion des clés. Les appariements doivent obligatoirement être réalisés de manière centralisée, avec les moyens informatiques et selon les processus de l'OFS. La remise et l'utilisation d'une clé dans le cas concret sont soumises à l'approbation de la direction et sont enregistrées.

Les droits d'accès aux différentes données (données d'input, données consolidées, données de production, voir chap. 7.1 et annexe 1) doivent être réglés clairement et de manière restrictive. Les différentes catégories de données doivent être cryptées séparément.

Les relevés et statistiques portant sur des données du degré de protection 3 doivent faire l'objet d'un règlement de traitement définissant précisément les mesures de sécurité prévues (art. 21, al. 1, let. a, OLPD). Un modèle a été établi sur la base des critères fixés par le PFPDT. Le règlement de traitement doit également préciser les appariements systématiques prévus dans le cadre de ces relevés et statistiques. Le modèle comprend un chapitre consacré aux appariements. Les règlements de traitement anciens sont à compléter d'un chapitre sur les appariements.

Tout appariement ne requiert pas l'établissement d'un règlement de traitement: le respect d'un règlement de traitement s'impose pour les appariements de données du degré de protection 3; dans le cas des données correspondant aux niveaux de protection 1 ou 2, un règlement de traitement ne devra être établi que si les données de l'output nécessitent un degré de protection 3. Si tel n'est pas le cas, il

Directives sur l'appariement

suffira d'appliquer les présentes directives en combinaison avec les dispositions légales pertinentes (notamment l'art. 13h ss de l'ordonnance sur les relevés statistiques et l'ordonnance du DFI sur l'appariement de données). Ces dispositions imposent cependant, au nom de la transparence, une documentation détaillée de tous les appariements. Les informations correspondantes figurent dans les formulaires de demande et dans la décision du directeur, conservés dans les archives électroniques de l'OFS.

Les données appariées peuvent être remises à des tiers aux conditions fixées à l'art. 9 de l'ordonnance sur les relevés statistiques, pour autant qu'elles soient destinées à des fins non personnelles (recherche, planification, statistique, etc.), comme le prévoit la loi. Cette remise dépend du degré de sensibilité des données, du but de leur utilisation et du destinataire. L'OFS n'est pas tenu de remettre des données à des tiers.

Les services statistiques cantonaux et communaux doivent garantir le même degré de protection des données et de respect du secret statistique que l'OFS, sans quoi ce dernier ne les autorise pas à appairer des données.

L'annexe 2 contient les lois pertinentes.

6 Bases de l'appariement

6.1 Définitions

6.1.1 Qu'est-ce qu'un appariement?

L'OFS applique la définition suivante: un appariement de données consiste à relier des **données individuelles** provenant de **diverses sources** (art. 13h de l'ordonnance sur les relevés statistiques) afin de produire un **nouveau jeu de données**. Chaque nouvelle exploitation d'un registre ou nouvelle réalisation d'une enquête en fait une nouvelle source de données². Un appariement au sens de la statistique publique implique qu'au **moins une des sources de données** soit élaborée **dans le cadre de la loi sur la statistique fédérale** et que l'appariement serve à **remplir des tâches statistiques** (art. 14a de la LSF).

On fait la distinction entre les catégories d'appariement suivantes (voir aussi le chap. 6.1.2 Catégories d'appariement)

- appariements systématiques destinés à la production statistique,
- appariements longitudinaux,
- appariements destinés à l'analyse statistique,
- cas spéciaux.

Les cas suivants ne constituent pas des appariements:

- Séries chronologiques de données agrégées (par ex. indices de prix, évolution de la production et de l'emploi, enquête sur la population active, etc.). Il ne s'agit en effet pas de données individuelles (voir aussi le ch. 6.1.2 → point 2 de la p. 9)
- Combinaison de données avec des nomenclatures ou ajout de géocoordonnées à l'aide d'adresses. Ces opérations servent uniquement à caractériser ou à structurer des informations.
- Regroupement de données de différentes sources dans le REE et dans le RegBL, tel que prévu dans les ordonnances respectives. Toute combinaison de données sortant de ce cadre est soumise aux prescriptions en matière d'appariement énoncées dans l'ordonnance sur les relevés statistiques, l'ordonnance du DFI sur l'appariement de données et le règlement de traitement correspondant.

² On peut citer comme exemple les appariements longitudinaux.

Directives sur l'appariement

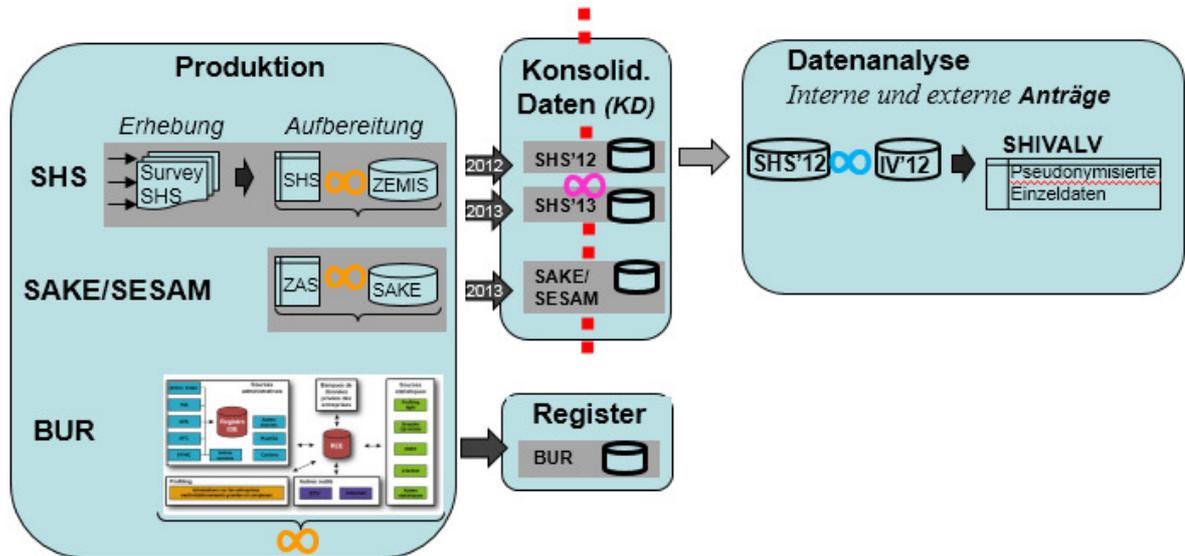
- Les enquêtes par panel, dans lesquelles les personnes répondent aux mêmes questions à différents intervalles. Ces enquêtes sont conçues au départ pour une réutilisation de l'échantillon et des réponses données lors de la précédente interview. Cette caractéristique est mentionnée dans l'ordonnance sur les relevés statistiques. On a ici affaire à une seule source de données³.
- Comparaison avec des données obtenues les années précédentes sur la base de la même enquête aux fins de plausibiliser les données et d'en vérifier la qualité. Les projets d'appariements visant à tester des méthodes complexes ont par contre valeur d'appariements et doivent être enregistrés comme tels (cas spéciaux).

6.1.2 Catégories d'appariements

Pour garantir une bonne acceptation des appariements et une pratique homogène en la matière, l'OFS a opté pour une structure d'appariement aujourd'hui bien établie parmi ses collaborateurs et les spécialistes de la statistique: *le processus de valorisation de la production statistique* (voir aussi le chap. 7.1). Chaque appariement doit être justifié et reposer sur une base légale. La plupart des activités statistiques sont énumérées dans *l'ordonnance sur les relevés statistiques*.

A partir de ce qui précède, on peut répartir les appariements en quatre catégories:

1. appariements systématiques destinés à la production statistique,
2. appariements longitudinaux,
3. appariements destinés à l'analyse statistique,
4. cas spéciaux (n'apparaissent pas sur le graphique).



1. Appariements systématiques destinés à la production statistique:

Les appariements systématiques⁴ qui entrent dans la production statistique de données consolidées (appariements de données d'input) sont énumérés dans l'ordonnance sur les relevés statistiques (SHS par ex.). Ils reposent tous sur une base légale, ce qui garantit la transparence. Peu importe qu'un relevé de données (primaires) ait été au préalable réalisé ou non à l'OFS. Parmi les statistiques de production mentionnées dans l'ordonnance, certaines s'appuient sur l'appariement de deux sources de données (par ex. ESPA et SESAM).

³ A l'inverse, la combinaison de données de l'ESPA et de la SILC avec des données de STATPOP et du REE a valeur d'appariement au sens des présentes directives.

⁴ Au sens de régulier.

Directives sur l'appariement

L'ordonnance ne mentionne pas les comparaisons effectuées avec d'autres sources de données uniquement à des fins qualitatives ou pour préparer des données consolidées destinées à la production.

Enrichissement de répertoires et de registres

Un processus particulier de la production statistique consiste à appairer des registres et des répertoires afin d'en enrichir le contenu. Ces recueils constituent eux aussi des sources de données administratives. De tels regroupements de données (REE, STATENT, RegBL, etc.) représentent aussi des appariements et doivent figurer dans l'ordonnance sur les relevés statistiques. Ces appariements présentent la particularité d'avoir leur propre base légale, tandis que les registres qu'ils alimentent sortent du cadre de valorisation statistique «classique».

Caractéristiques des appariements systématiques destinés à la production statistique

- Ils sont internes à l'OFS.
- Ils visent la production statistique.
- Ils figurent dans l'ordonnance sur les relevés statistiques (comme relevé ou comme statistique).
- Ils nécessitent une autorisation unique. Si l'appariement figure dans l'annexe de l'ordonnance, l'autorisation d'appariement est donnée dans le cadre de la révision annuelle de cette dernière.

Enrichissement de répertoires et de registres

- Ils reposent sur leur propre ordonnance.
- Les sources qui sont reliées aux répertoires et aux registres sont clairement indiquées dans l'ordonnance sur les relevés statistiques.

2. Appariements longitudinaux

Les appariements longitudinaux consistent à associer à des moments différents (au minimum une fois par année) les variables d'une même unité présentes dans des enquêtes ou des sources de données administratives identiques afin d'identifier des changements et des évolutions. Les appariements concernent uniquement des données individuelles et pas les séries temporelles obtenues à partir de données agrégées (p. ex. indices de prix, évolution de la production, taux de sans-emploi, etc.).

Exemples positifs d'appariements longitudinaux:

- 1) Combinaison de données individuelles tirées de relevés réalisés à des moments différents pour représenter une évolution (exemple: démographie des entreprises).
- 2) Etablissement de parcours de formation individuels par la combinaison d'informations tirées de relevés qui se rapportent à des périodes différentes.

Exemples négatifs d'appariements longitudinaux:

- 1) Comparaison avec des données individuelles de l'année précédente afin de contrôler les réponses et la qualité (exemples: statistique de la valeur ajoutée, statistique des caisses de pensions).
- 2) Enquêtes par panel, dans lesquelles les personnes répondent aux mêmes questions à différents moments (exemples: SILC, ESPA). Ces enquêtes sont conçues au départ pour une réutilisation de l'échantillon et des réponses données lors de la précédente interview. Cette caractéristique est mentionnée dans l'ordonnance sur les relevés statistiques.

Les appariements longitudinaux sont effectués par l'OFS et utilisés à l'interne. Ils consistent normalement à relier les données consolidées d'une même enquête réalisée à différentes périodes. Le résultat de cet appariement peut être considéré comme un produit figurant dans l'inventaire de l'OFS. Les utilisateurs externes qui souhaitent recourir à de telles données consolidées sur plusieurs unités temporelles peuvent le faire moyennant la conclusion d'un contrat de protection des données ordinaire. Un tel contrat est aussi nécessaire lorsque la livraison des données n'englobe pas toutes les variables du produit.

Directives sur l'appariement

Caractéristiques des appariements longitudinaux

- essentiellement à des fins internes (il est possible de réaliser des appariements longitudinaux pour des externes, mais ils tombent dans ce cas dans la catégorie des «appariements destinés à l'analyse statistique»)
- destinés en premier lieu à l'analyse
- destinés généralement à des «produits standardisés»
- appariements internes: procédure d'autorisation automatique (ajout dans le règlement de traitement du relevé)
- inscription à l'inventaire des appariements longitudinaux de l'OFS
- l'inscription dans l'ordonnance sur les relevés statistiques est obligatoire pour les appariements systématiques (cf. art. 13n de l'ordonnance)

Les externes peuvent obtenir les données appariées dans le cadre d'un contrat de protection des données.

3. Appariement destinés à l'analyse statistique

Les appariements statistiques (destinés à l'output) que l'OFS effectue en vertu de l'art. 14a LSF pour remplir ses tâches statistiques doivent avoir fait l'objet d'une demande documentée pour des raisons de légalité et de transparence. Cette règle vaut aussi bien pour les demandes internes que pour les demandes externes. Ces appariements ne servent pas à la production, mais à l'analyse de données (exemples: parcours de formation, projet NCCR de l'Université de Genève).

Caractéristiques des appariements destinés à l'analyse statistique

- utilisation de données appariées à des fins internes ou externes
- destinés à des fins d'analyse
- généralement uniques
- demandes fondée sur le processus décrit au chap. 0 et dans l'annexe 5
- inscription dans l'inventaire de l'OFS «appariements destinés à l'analyse»

4. Cas spéciaux

Les appariements qui n'appartiennent pas à l'une des trois catégories citées (à savoir tous les autres) sont également clairement mentionnés et sont regroupés dans la catégorie des cas spéciaux. De manière empirique, on peut dire que les projets d'appariement méthodologiquement complexes, comme les tests de faisabilité techniques, sont à considérer comme des cas spéciaux, mais que cette catégorie exclut les comparaisons avec des données de l'année précédente aux fins de plausibilisation et de contrôle de la qualité.

Caractéristiques des cas spéciaux

- si internes: uniques pour les tests de faisabilité technique
- demande à formuler selon le processus décrit au chap. 6.3 des présentes directives
- inscription dans l'inventaire de l'OFS «appariements pour cas spéciaux»

6.2 Principes applicables

Les expériences ont confirmé la nécessité d'instaurer une approche intégrale et de définir l'appariement de manière suffisamment large. Cela passe par la définition de principes généraux qui soient valables pour tout l'OFS et que les unités opérationnelles puissent appliquer selon leurs spécificités.

L'appariement des données à l'OFS se fonde sur les principes suivants:

- L'OFS encourage et met en œuvre de manière ciblée l'appariement de données pour obtenir des informations statistiques de manière efficiente, en préservant les ressources.
- On ne procède à des appariements qu'à des fins statistiques au sens de la LSF; les appariements doivent être nécessaires et répondre au principe de proportionnalité. Le besoin d'information doit être avéré; les appariements ne sont pas faits pour constituer des réserves de données.
- Les appariements sont effectués en toute transparence et sont bien documentés. Leur base légale est vérifiée et ils sont inventoriés puis approuvés par la direction de l'OFS.
- Les appariements doivent se justifier et être effectués selon une méthode correcte. Les données d'input comme les données d'output doivent satisfaire aux exigences de qualité statistique usuelles.
- Les processus d'appariement sont soumis à des conditions strictes en matière de protection et de sécurité des données. Ils sont contrôlés et documentés de manière centralisée.
- Leurs différentes étapes doivent être aménagées aussi simplement que possible. On évitera autant que possible les complications inutiles et les obstacles administratifs.
- Les étapes des processus représentés dans ces directives (cf. chap. 6.3 Prescriptions relatives aux appariements, chap. 0 Processus d'appariement et annexe 5: Description des processus) s'appliquent à tous les mandants (internes et externes).
- Les appariements de données au sens de l'art. 14a LSF sont effectués principalement par l'OFS. Si les conditions sont remplies, certaines tâches d'appariement peuvent être déléguées à des tiers (voir 10.3.2).
- L'OFS vérifie que les conditions légales sont remplies (exemples: le projet proposé repose-t-il sur des bases légales et est-il conforme à ces dernières?) et que l'appariement est possible sur les plans méthodologique, technique et du contenu si possible dans les 15 jours ouvrables suivant le dépôt de la demande.
- La transmission de données appariées est possible dans le cadre de la réglementation applicable à la transmission de données individuelles (art. 13I de l'ordonnance sur les relevés statistiques). Elle dépend du degré de sensibilité des données, du but de leur utilisation et du destinataire. Un contrat de protection des données doit être conclu dans tous les cas.
- Les dispositions applicables comprennent, outre les règles exposées dans le présent document, les prescriptions définies dans d'autres guides de l'OFS au sujet du traitement et de la remise de données individuelles ainsi que de la protection et de la sécurité des données⁵.
- Le coût des appariements est calculé d'après les barèmes de l'ordonnance sur les émoluments et indemnités perçus pour les prestations de services statistiques des unités administratives de la Confédération.

⁵ Directives relatives à la protection des données en cas de transmission de données individuelles à des tiers, 18.10.2006



6.3 Prescriptions relatives à l'appariement

Catégories et conditions de l'appariement: Interne

	Conditions à remplir			
Catégorie de l'appariement	Avant	Déroulement (voir aussi le chap. 10.3.1)	Après	Réserves/facilitations
1) Appariements systématiques destinés à la production statistique	<ul style="list-style-type: none"> • Demande selon formulaire (cf. annexe 5) • Avis SPROD, METH et RD • Décision CD, information du reste de la GL • Décision de la direction • SPROD demande à IT d'établir la clé • La partie du règlement de traitement concernant l'appariement est mise à jour (uniquement pour le degré de protection 3) • Demande d'inscription dans l'annexe de l'ordonnance sur les relevés statistiques • Inscription dans l'inventaire 	<ul style="list-style-type: none"> • Remise de la clé au responsable de secteur • Dé-pseudonymisation (du degré 2 au degré 1, cf. chap. 7.3) • Pseudonymisation/appariement spécifique au cas 	<ul style="list-style-type: none"> • Fin de la validité de l'autorisation d'utiliser la clé 	<ul style="list-style-type: none"> • Aucune demande ne doit être déposée pour les appariements qui sont déjà mentionnée dans l'annexe de l'ordonnance sur les relevés statistiques. • Pas d'effacement du jeu de données apparié

Directives sur l'appariement

Enrichissement de répertoires et de registres	<ul style="list-style-type: none"> • Documentation • Règlement de traitement (seulement pour le degré de protection 3) • Base légale spéciale 		<ul style="list-style-type: none"> • Documentation finale 	<ul style="list-style-type: none"> • Possibilité de simplifier la procédure pour les données non sensibles • Pas d'effacement du jeu de données apparié
2) Appariements longitudinaux	<ul style="list-style-type: none"> • Demande selon formulaire (cf. annexe 6) • Avis SPROD, METH et RD • Décision CD, information du reste de la GL • Décision de la direction • SPROD demande à IT d'établir la clé • La partie du règlement de traitement concernant l'appariement est mise à jour (uniquement pour le degré de protection 3) • Demande d'inscription dans l'annexe de l'ordonnance sur la statistique • Inscription dans l'inventaire 	<ul style="list-style-type: none"> • Remise de la clé au responsable de secteur • Dé-pseudonymisation (du degré 2 au degré 1, cf. chap. 7.3) • Pseudonymisation/appariement spécifique au cas 	<ul style="list-style-type: none"> • Fin de la validité de l'autorisation d'utiliser la clé • 	<ul style="list-style-type: none"> • Possibilité de simplifier la procédure pour les données non sensibles • Pas d'effacement du jeu de données apparié • Mesures de sécurités supplémentaires ou limitations dans la remise et/ou l'utilisation des données sensibles
3) Appariements destinés à l'analyse statistique,	<ul style="list-style-type: none"> • Demande selon formulaire (cf. annexe 6) • Avis SPROD, METH et RD • Décision CD, information du reste de la GL • Décision de la direction • SPROD demande à IT d'établir la clé • Elaboration d'un règlement de traitement spécial lors le degré de protection des données d'output passe à 3 	<ul style="list-style-type: none"> • Remise de la clé au responsable de secteur • Dé-pseudonymisation (du degré 2 au degré 1, cf. chap. 7.3) • Pseudonymisation/appariement spécifique au cas 	<ul style="list-style-type: none"> • Fin de la validité de l'autorisation d'utiliser la clé • 	<ul style="list-style-type: none"> • Possibilité de simplifier la procédure pour les données non sensibles • Mesures de sécurités supplémentaires ou limitations dans la remise et/ou l'utilisation des données sensibles • Effacement des appariements de données sensibles une fois l'analyse terminée

Directives sur l'appariement

	<ul style="list-style-type: none"> • Inscription dans l'inventaire 			
4) Cas spéciaux (appariements tests) à des fins internes	<ul style="list-style-type: none"> • Processus de demande simplifié par la voie hiérarchique (sans RD/METH) • Appréciation par la section concernée • Décision CD, information du reste de la GL • Décision de la direction • SPROD demande à IT d'établir la clé • La partie du règlement de traitement concernant l'appariement est mise à jour (uniquement pour le degré de protection 3) • Inscription dans l'inventaire 	<ul style="list-style-type: none"> • Remise de la clé au responsable de secteur • Dé-pseudonymisation (du degré 2 au degré 1, cf. chap. 7.3) • Pseudonymisation/appariement spécifique au cas 	<ul style="list-style-type: none"> • Fin de la validité de l'autorisation d'utiliser la clé • 	<ul style="list-style-type: none"> • Inscription à l'inventaire de l'OFS • Possibilité de prévoir des mesures de sécurités supplémentaires pour l'appariement de données sensibles • Effacement des appariements de données sensibles une fois les tests terminés

Catégories et conditions de l'appariement: domaine externe

Catégorie de l'appariement	Conditions requises			
	Avant	Exécution	Après	Réserves/facilitations
1) Appariements systématiques destinés à la production statistique,	-	-	-	-
Enrichissement de répertoires et de registres	-	-	-	-
2) Appariements longitudinaux	<ul style="list-style-type: none"> • Demande par formulaire • Avis SPROD, METH et RD • Décision CD, information du reste de la GL • Décision de la direction 	<ul style="list-style-type: none"> • Remise de la clé au responsable de secteur • Dé-pseudonymisation (du degré 2 au degré 1, cf. chap. 7.3) 	<ul style="list-style-type: none"> • Fin de la validité de l'autorisation d'utiliser la clé 	<ul style="list-style-type: none"> • Possibilité de simplifier la procédure pour les données non sensibles • Mesures de sécurités supplémentaires ou limitations dans

Directives sur l'appariement

	<ul style="list-style-type: none"> • SPROD demande à IT d'établir la clé • Etablissement d'un contrat de protection et d'appariement des données • Inscription dans l'inventaire 	<ul style="list-style-type: none"> • Pseudonymisation/appariement spécifique au cas • Event. examen des données d'output par l'OFS • 		<p>la remise et/ou l'utilisation des données sensibles</p> <ul style="list-style-type: none"> • Des tiers peuvent au besoin être associés aux travaux d'appariement
3) Appariements destinés à l'analyse statistique,	<ul style="list-style-type: none"> • Demande par formulaire • Avis SPROD, METH et RD • Décision CD, information du reste de la GL • Décision de la direction • Etablissement d'un contrat de protection et d'appariement des données • Inscription dans l'inventaire 	<ul style="list-style-type: none"> • Remise de la clé au responsable de secteur • Dé-pseudonymisation (du degré 2 au degré 1, cf. chap. 7.3) • Appariement par/à l'OFS • Event. examen des données d'output par l'OFS • Pseudonymisation • Anonymisation (output) 	<ul style="list-style-type: none"> • • Fin de la validité de l'autorisation d'utiliser la clé-communication des données 	<ul style="list-style-type: none"> • Possibilité de simplifier la procédure pour les données non sensibles • Mesures de sécurités supplémentaires ou limitations dans la remise et/ou l'utilisation des données sensibles • Effacement des appariements de données sensibles une fois l'analyse terminée • Des tiers peuvent au besoin être associés aux travaux d'appariement (voir 10.3.2)
4)		-	-	-

- La validation de la clé est soumise à la décision de la direction concernant la réalisation d'un appariement. La section IT gère du point de vue technique l'outil pour générer et utiliser la clé, sur mandat du service spécialisé en appariements. Cela implique l'existence d'un numéro de référence GEVER, car ce dernier est lié à la décision de la direction.

4 Sensibilité des données

Lors des appariements, il importe d'accorder une grande importance à la sensibilité des données, qu'il s'agisse des données qui sont appariées (données d'input) ou des données qui résultent de l'appariement (données d'output). Cette sensibilité détermine le caractère plus ou moins strict des mesures à prendre (restriction de la remise des données, effacement, etc.). Plus les données sont sensibles, plus leur protection doit être élevée. On distingue quatre degrés de protection:

Degrés de protection

(tiré des "Directives relatives à la protection des données en cas de transmission de données individuelles à des tiers", état au 18.10.2006)

- **Degré 0 (données techniques):**
Données ne se rapportant pas à une personne, par ex. résultats de mesures.
- **Degré 1 (données personnelles simples):**
Données dont la communication ne risque pas de porter atteinte à la personnalité des personnes physiques ou morales concernées (par ex. le nom, le prénom, l'adresse, la date de naissance, la branche de l'entreprise). Combinées à des données sensibles, ces données peuvent toutefois requérir un degré de protection plus élevé (par ex. les noms des détenus d'un établissement pénitentiaire ou de patients d'une clinique traitant des malades du sida). Souvent, les données du degré 1 sont assez facilement accessibles (via l'annuaire téléphonique, les rapports annuels et d'autres publications) et peuvent aussi être relevées par des tiers, indépendamment de l'OFS, moyennant certains travaux.
- **Degré 2 (données personnelles qualifiées):**
Données dont la communication présente un certain risque de porter atteinte à la personnalité des personnes physiques et morales concernées (par ex. les données n'appartenant pas au degré 3 qui portent sur les revenus et la fortune, le loyer, les relations commerciales, la formation, l'activité professionnelle, les opinions et les comportements).
- **Degré 3 (données personnelles sensibles):**
Données dont la communication présente un risque majeur de porter atteinte à la personnalité des personnes physiques ou morales concernées (par ex. indications sur la religion, sur les opinions et activités idéologiques, syndicales et politiques, la race, la santé, la sphère privée, l'aide sociale, les infractions).

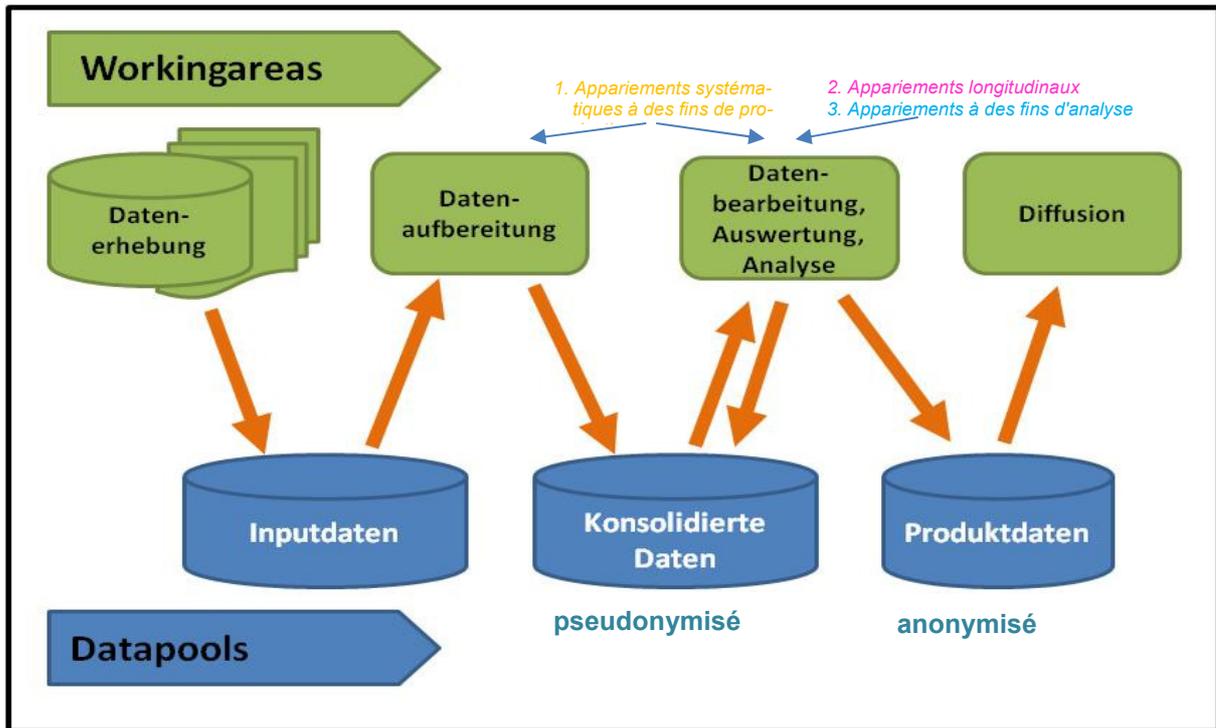
7 Production statistique, sécurité et appariement des données

7.1 Processus de production statistique

Dans la production statistique, on distingue généralement les étapes suivantes:

⇒ relevé/saisie des données ⇒ préparation des données ⇒ traitement des données, exploitation, analyse ⇒ diffusion.

Le schéma ci-après s'applique normalement à l'ensemble de la production statistique.



Voici comment se décomposent ces étapes de la production statistique

Relevé et saisie des données

- Collecte des données au moyen de relevés (*directs ou fondés sur des registres, notamment administratifs*) et d'enquêtes bien documentés (bases conceptuelles, descriptifs d'interface pour les variables, etc.), ou reprise des données de sources administratives, de mesures, etc.
- Saisie des données.
- Enregistrement des données d'input (*les données d'input sont les données brutes provenant des relevés. Elles contiennent normalement les caractères d'identification des personnes qui sont nécessaires au dépouillement et aux demandes de renseignement complémentaires -> voir l'annexe 1.*)

Préparation des données

- Les données d'input sont préparées par l'élimination des incohérences et le traitement des valeurs manquantes ou extrêmes, et des variables de synthèse (caractères dérivés) sont calculées. Des appariements sont possibles à ce stade. L'annexe de l'ordonnance sur les relevés mentionne les appariements de données systématiques de la production statistique, qui font l'objet d'une documentation détaillée (conceptions générale et détaillée, règlement de traitement, SIPD, etc.) (**appariements destinés à la production statistique**).
- Les données préparées sont stockées sous la forme de «données consolidées» (*les données consolidées sont des données individuelles qui ont été apurées. Elles sont conservées sous une forme*

Directives sur l'appariement

pseudonymisée. Les variables d'identification (nom, adresse, etc.) sont supprimées et les identificateurs (numéro AVS, IDE, etc.) sont pseudonymisés (voir l'annexe 1). Les bonnes pratiques commandent de stocker les données consolidées dans la banque de données centrale qui leur est destinée, conformément au concept SIS.

L'accès aux appariements a lieu uniquement par des processus ETL sécurisés. Il n'est pas possible d'accéder directement aux données.

Traitement, exploitation et analyse des données

- Les données consolidées font l'objet d'autres opérations, telles que des opérations d'exploitation, d'analyse ou d'appariement (**appariements longitudinaux** ou **appariements à des fins d'analyse**).
- Ces diverses opérations sont en principe pseudonymisées.
- Les données qui en résultent ont valeur de «données de production» (*qui se présentent généralement sous une forme agrégée, et donc anonymisée -> voir l'annexe 1*) ou rejoignent les autres «données consolidées» si elles ne doivent pas être effacées en raison de leur caractère sensible (cf. art. 14a LSF).

Diffusion

- A partir des données de production, on crée ensuite des articles qui sont diffusés sous une forme appropriée aux clients intéressés.
- Les données consolidées forment la base des produits de l'OFS (tableaux, cubes, comptoirs de données pour des contrats de livraison des données ou des cartes, etc.).

7.2 Protection des données

L'OFS garantit la protection et la sécurité des données également lors des appariements. Ces derniers retiennent de plus en plus l'attention du public. Cela oblige l'OFS à pratiquer la transparence sur ses activités, afin de préserver la confiance placée en lui.

Les collaborateurs de l'OFS, qui connaissent les avantages des appariements, mais aussi les risques qu'ils représentent, sont en mesure d'informer de manière claire le public sur ses activités. Ils sont capables de répondre rapidement et de manière transparente et détaillée aux demandes d'information sur l'activité statistique qui leur parviennent d'autres offices ou du PFPDT.

L'OFS est tenu d'appliquer la loi sur la statistique fédérale, la loi sur la protection des données, le code de bonnes pratiques, la charte de la statistique publique et les directives concernant la sécurité informatique dans l'administration fédérale. Chaque collaborateur respecte les mesures prescrites par ces textes, par ex. en s'engageant par signature à respecter les règles de bonne pratique. Les divisions définissent les secteurs, puis les responsables des secteurs définissent les règles d'accès aux données d'input et aux sources de données utilisables par le secteur. Les secteurs garantissent la pseudonymisation des données consolidées et leurs responsables veillent à ce que seuls les ayant-droits puissent y accéder. Le règlement des rôles défini au chap. 8 précise les détails.

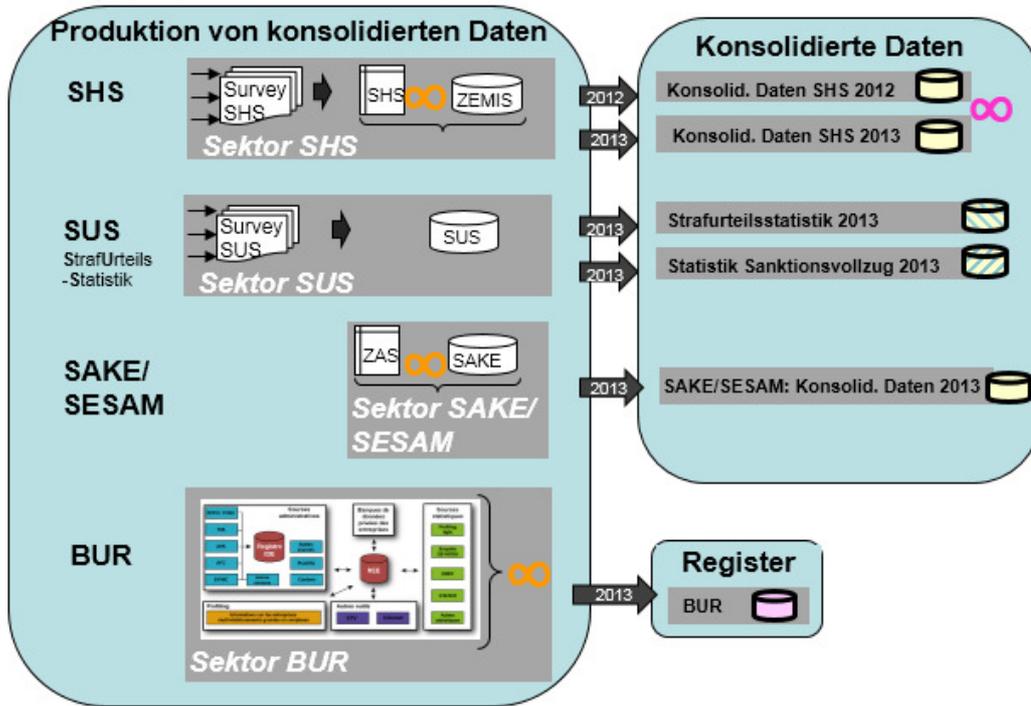
Il existe trois niveaux de mesures destinées à protéger les données appariées d'un accès non autorisé (la pseudonymisation et la gestion des clés offrent des sécurités supplémentaires):

7.2.1 Espace de confiance à l'OFS

L'espace de confiance à l'OFS est défini par les lois pertinentes, la charte de la statistique publique et les codes de bonne pratique. Ces textes, que tous les collaborateurs de l'OFS sont tenus d'appliquer, constituent le premier niveau de protection.

7.2.2 Secteurs

Le niveau des secteurs correspond aux unités de production à l'OFS. Leur but est la production, à partir des données collectées, de données consolidées destinées à être traitées, exploitées et analysées. Le graphique suivant livre l'exemple de quatre secteurs: SHS, SUS, ESPA/SESAM et REE.



Chaque secteur dispose normalement d'une application de relevé dans laquelle sont enregistrées les données d'input, comme dans le cas de la SHS ou de la SUS. On peut aussi qualifier de secteur les unités de production qui ne relèvent pas de données d'input mais qui réalisent des produits statistiques à partir de différentes sources (par ex. ESPA/SESAM).

Les secteurs protègent les données en limitant et en contrôlant l'accès aux données d'input. Seuls les collaborateurs d'une unité de production autorisée peuvent accéder à l'application de relevé (SHS) ou, en l'absence de cette dernière, à l'outil de production (ESPA/SESAM)

Plus la sectorisation est fine, plus la possibilité d'apparier facilement des données est limitée.

7.2.3 Pseudonymisation des données consolidées

La pseudonymisation des données consolidées empêche l'appariement non autorisé de données individuelles consolidées (les pseudo-identificateurs sont propres à chaque processus de consolidation et ne sont pas compatibles avec les autres jeux de données consolidées). Les sections règlent l'attribution et la gestion des droits d'accès. L'accès aux données consolidées est lui aussi réglementé et contrôlé; il est réservé aux collaborateurs autorisés de l'OFS.

En ce qui concerne les relevés de données consolidées qui sont effectués dans des intervalles de plusieurs années, de nouveaux pseudoidentificateurs sont créés à chaque relevé. Pseudonymisation et dé-pseudonymisation

Pseudonymisation

La pseudonymisation consiste à supprimer tous les caractères personnels d'un jeu de données qui permettent d'identifier la personne à laquelle il se rapporte (par ex. nom, adresse) et de les remplacer par un pseudonyme créé par l'OFS. Les données pseudonymisées ne permettent plus de remonter à

Directives sur l'appariement

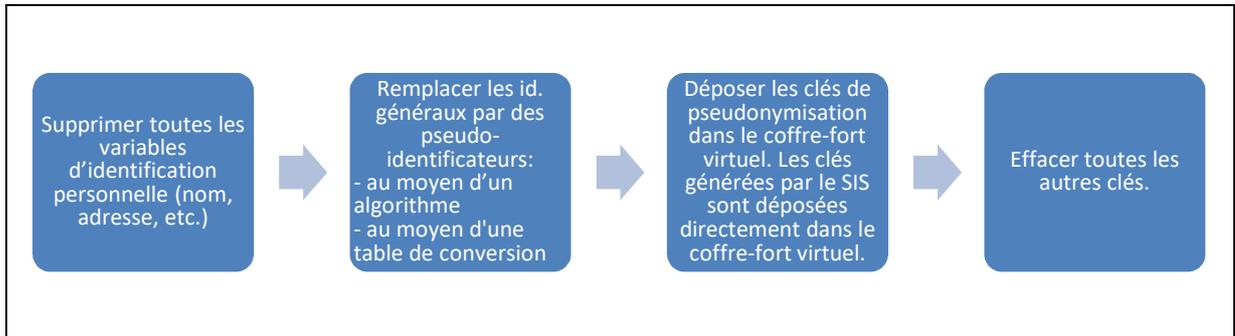
la personne à laquelle elles se rattachent.

Les pseudoidentificateurs sont soit générés par un algorithme de cryptage (au sein du SIS), soit attribués directement (attribution de nouveaux identificateurs et création d'un tableau de conversion). Les données qui se trouvent à l'extérieur du SIS doivent aussi être pseudonymisées au moyen du programme de cryptage. La pseudonymisation s'effectue au sein du SIS par la transformation de l'identificateur d'origine en un pseudoidentificateur, au moyen d'un processus à deux niveaux (niveau 1: clé OFS, niveau 2: clé sectorielle)

Des clés et des pseudoidentificateurs spécifiques sont générés pour chaque production (relevé) de données consolidées. Cette méthode permet d'éviter que des jeux de données différents concernant une même personne puissent être appariés sans information supplémentaire.

Les clés utilisées pour la pseudonymisation (algorithmes, tableaux de conversion) sont conservées dans un répertoire sécurisé spécial (selon le principe du coffre-fort virtuel; voir le chap.7.3) et supprimées partout ailleurs. Les clés sont alors accessibles uniquement via le coffre-fort virtuel. Les clés sont générées et administrées par la section IT sur mandat du service spécialisé en appariements.

→ Processus de pseudonymisation

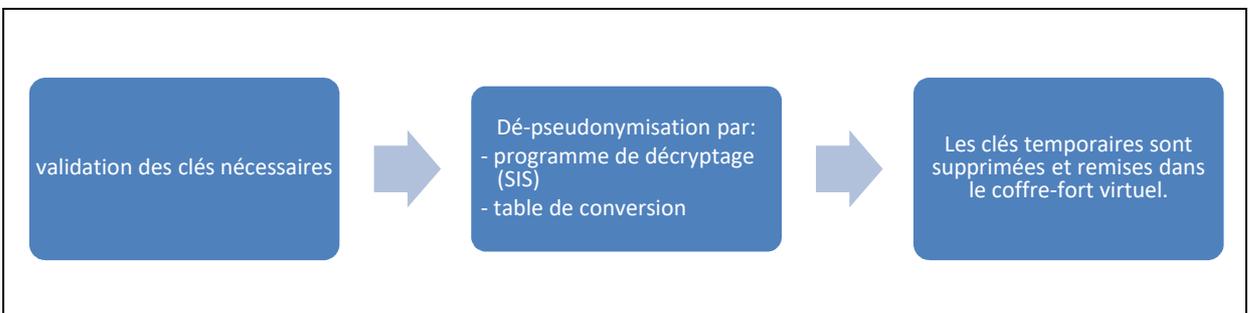


Dé-pseudonymisation

Le processus inverse est la dé-pseudonymisation. Pour pouvoir appairer les données de différents relevés/secteurs, il faut d'abord les dé-pseudonomiser, à savoir décrypter les identificateurs, si les données ne proviennent pas du même secteur. Lorsque l'appariement concerne des données de l'OFS, la dé-pseudonymisation ne comporte qu'une étape (du niveau 2 au niveau 1). Les identificateurs d'origine sont reconstitués, puis remplacés par un identificateur spécifique au projet, grâce auquel les données concernées peuvent être appariées.

- La section IT remet la clé tirée du coffre-fort virtuel au responsable du secteur.
- Le programme de décryptage est lancé.
- Grâce à la clé, les pseudo-identificateurs sont remplacés par les identificateurs d'origine.
- Le responsable de secteur établit l'identificateur spécifique au projet pour l'appariement

→ Processus de dé-pseudonymisation



7.3 Principe de la gestion des clés

Seuls les détenteurs des clés de pseudonymisation et des droits d'accès correspondants peuvent appairer des données. Le contrôle des appariements est garanti par la conservation des clés dans un fichier protégé (coffre-fort virtuel) et par une gestion centralisée des clés par les seules personnes autorisées.

Conception:

Les clés sont conservées de manière centralisée et sécurisée dans un **coffre-fort virtuel**. Cet outil informatisé est présent dans tout l'OFS et permet à la section IT de gérer les clés et d'octroyer les autorisations de les utiliser. C'est la direction qui détermine qui a accès au coffre-fort virtuel (voir le chap. 8) et l'accès est implicitement garanti par l'autorisation de procéder à l'appariement dans le processus GEVER.

C'est la section IT qui gère le coffre-fort virtuel central, les clés qu'il contient et les accès à ces dernières. Elle octroie au responsable sectoriel concerné le droit d'utilisation et d'accès pour la dé-pseudonymisation et la création des clés. Le coffre-fort virtuel correspond à la solution de l'environnement informatique développée par la section IT. L'accès au coffre-fort est sécurisé et la section IT et le responsable sectoriel concerné ne peuvent y accéder que s'ils disposent des codes attribués uniquement pour les processus ETL requis.

Pour des raisons liées aux techniques de sécurisation, seul le responsable sectoriel disposant de l'autorisation du service spécialisé en appariements via la section IT peut dé-pseudonymiser des fichiers de son secteur.

La section IT est aussi responsable de la gestion de la clé sectorielle créée pour le jeu de données résultant de l'appariement. Un principe essentiel veut que la section IT n'ait pas accès aux données statistiques.

Une documentation concernant la pseudonymisation/infrastructure informatique (disponible uniquement à l'intérieur de l'OFS) est enregistrée sous <https://intranet.confluence.bfs.ad-min.ch/pages/viewpage.action?pageId=27754683>

8 Organisation et rôles

Direction:

L'art. 2, al. 3, de l'ordonnance du DFI concernant la réalisation d'appariements de données statistiques octroie à la direction de l'OFS la responsabilité des appariements. Celle-ci est assumée concrètement par le directeur de l'OFS, qui peut consulter le conseil de direction. La direction décide de la réalisation des appariements et fixe les éventuelles réserves et conditions. L'autorisation d'effectuer un appariement vaut également autorisation d'utiliser les clés requises à cet effet.

Groupe de travail Appariements

Le groupe de travail Appariement est dirigé par un membre du conseil de direction. Il compte au moins un représentant de chaque division, du service juridique, de la section Méthodes statistiques, des services informatiques et du service d'appariement. Il se réunit en cas de nécessité, mais au moins deux fois par an.

Le mandat du groupe de travail a été attribué par le conseil de direction le 9 décembre 2015 et il prévoit les tâches suivantes :

Soutien, discussion et évaluation des travaux dans le cadre des appariements de données, avec les priorités thématiques suivantes :

- Orientation générale et conceptuelle des appariements de données et monitoring des travaux
- Mise en œuvre des opérations, y compris aux niveaux informatique (technique, contenu) et organisationnel
- Bases légales, notamment la protection des données et les conditions méthodologiques
- Communication et transparence (à l'intérieur et à l'extérieur de l'office)

Directives sur l'appariement

Service spécialisé en appariements

Le service spécialisé en appariements est responsable des activités suivantes:

- accompagner administrativement et coordonner les appariements de données. Cette activité inclut notamment le contrôle formel des demandes, l'élaboration et la transmission des avis internes (sous la forme d'une brève recommandation), la gestion des processus de décision, le pilotage, le contrôle formel et la documentation des projets d'appariement ainsi que l'octroi de renseignements. Le service d'appariement n'a aucun accès aux données à appairer.
- Débattre et émettre des recommandations, en cas de requête spéciale, au sujet du contenu, de la méthode, de la qualité, de la faisabilité et de l'éthique. Au besoin. Il peut faire appel à d'autres services (section Méthodes statistiques, service juridique, groupe de travail Appariements, etc.).
- gérer l'inventaire.

Responsable de secteur⁶:

Le responsable de secteur, chargé d'un secteur spécifique, assume les responsabilités suivantes:

- Il gère l'accès au secteur par les personnes autorisées.
- Il réalise la pseudonymisation et la dé-pseudonymisation, grâce aux droits que lui octroie le service spécialisé en appariements. Il est responsable de l'utilisation des clés pour son secteur. C'est la seule personne qui peut demander à la section IT d'autoriser l'utilisation des clés.
- Il assure la pseudonymisation des données consolidées et garantit que seules les personnes autorisées puissent accéder aux données consolidées.

Le rôle de responsable de secteur peut être confié à plusieurs membres d'une même unité organisationnelle, par exemple pour garantir une suppléance.

Collaborateur au secteur:

Le collaborateur au secteur est un collaborateur spécialiste du secteur. Il reçoit du responsable de secteur les droits d'accéder aux données consolidées du secteur.

Le rôle de collaborateur au secteur peut être confié à plusieurs membres d'une même unité organisationnelle, par exemple pour garantir une suppléance ou permettre l'accomplissement de tâches communes.

Autres rôles:

- Chef/fe de division: elle est la première instance de décision pour toutes les demandes d'appariement internes et externes. Elle garantit l'expertise requise et la mise à disposition des ressources nécessaires à l'appariement.
- Service juridique: il examine la conformité juridique des demandes d'appariement internes et externes.
- Section Méthodes statistiques: elle examine si les demandes d'appariement internes et externes sont réalisables et correctes sur le plan de la méthode statistique.
- Section IT : s'occupe, sur mandat du service spécialisé en appariements, de la gestion des clés et notamment de l'administration des accès au coffre-fort virtuel. La gestion des clés est assurée par un nombre limité de membres de la section IT. Ces personnes sont habilitées, sur la base d'une décision de la direction, à rendre les clés nécessaires à l'appariement accessibles aux responsables de secteur pour la de-pseudonymisation ou à attribuer les droits d'accès correspondants. Le service de gestion des clés veille en outre à une utilisation correcte des clés et la limitation des autorisations d'utiliser les clés au terme des appariements. La section IT n'a pas accès aux données individuelles.

⁶ Dans ce qui suit, le terme «responsable de secteur» ou «collaborateur au secteur» désigne un rôle et les droits d'accès qui y sont liés. Une même personne peut assumer différents rôles.

9 Éléments de l'appariement de données

9.1 Inventaires des appariements à l'OFS

Il existe quatre catégories d'appariement à l'OFS (voir chap. 6.1.2). Les inventaires donnent une vue d'ensemble de tous les appariements effectués (). La transparence et la protection des données sont ainsi garanties.. Le chap. 6.3 indique les conditions auxquelles chaque type d'appariement doit satisfaire.

Appariements destinés à la production statistique

Les appariements destinés à la production statistique sont indiqués comme tels dans l'ordonnance sur les relevés. L'inventaire qui y figure est de notoriété publique. L'indication des appariements destinés à la production statistique a été ajoutée à l'ordonnance en 2014, à l'occasion de sa révision annuelle.

- Publication: les relevés et les statistiques figurent aussi bien dans l'annexe de l'ordonnance sur les relevés statistiques que dans l'inventaire des appariements destinés à la production statistique. On peut y accéder par la page d'accueil du site de l'OFS.
- Responsable de la réalisation: service juridique
- Méthode: révision annuelle de l'annexe de l'ordonnance. Les unités spécialisées (secteurs) sollicitent des appariements de manière continue. La marche à suivre est régie dans le présent document (voir chap. 10 et annexe 5).
- Autorisation: la direction, puis le Conseil fédéral, décident de l'entrée en vigueur de l'ordonnance révisée.

Appariements longitudinaux

Les appariements longitudinaux sont normalement moins problématiques que les appariements transversaux, du fait que les données proviennent du même relevé réalisé à des moments différents. Le risque qu'on puisse en dégager des profils sensibles n'en existe pas moins. Les appariements longitudinaux doivent être eux aussi dûment inventoriés et soigneusement réalisés.

- Publication: inventaire des appariements de données
- Responsable de la mise à jour permanente: service d'appariement
- Méthode: les secteurs adressent au service d'appariement des demandes régulières d'appariement. La marche à suivre est régie dans le présent document (voir chap. 10 et annexe 5)
- Autorisation: par le directeur (qui peut consulter le conseil de direction)

Appariements destinés à l'analyse de données

Normalement, les appariements destinés à l'analyse de données ne sont pas systématiques (c'est-à-dire qu'ils ne sont pas récurrents). Les demandes d'appariement de données peuvent émaner de l'OFS ou de tiers. L'ordonnance du DFI sur l'appariement de données (voir annexe 2, point 4) règle leur documentation par le service d'appariement.

- Publication: inventaire des appariements de données
- Responsable de la mise à jour permanente: Service d'appariement
- Méthode: les demandes sont déposées au fur et à mesure et traitées par le service d'appariement, selon le processus prévu dans le présent document (voir chap. 10 et annexe 5)
- Autorisation: par le directeur (qui peut consulter le conseil de direction)

Cas spéciaux

Tous les appariements qui ne figurent pas dans l'une des trois catégories ci-dessus sont également inventoriés. Cela permet de recenser la totalité des appariements.

- Publication: inventaire des cas spéciaux d'appariement, non public.
- Responsable de la mise à jour permanente: service d'appariement
- Méthode: demandes déposées au fur et à mesure au service spécialisé en appariements.
- Autorisation: par le directeur (qui peut consulter le conseil de direction).

Directives sur l'appariement

9.2 Demandes d'appariement

Les demandes d'appariement doivent être déposées par écrit (voir annexe 6), normalement auprès du service spécialisé en appariements ou de la section de production concernée. Lorsqu'une telle demande est adressée à une section de production, il faut y associer immédiatement le service spécialisé en appariements.

Données requises

Le service spécialisé en appariements examine que la documentation accompagnant la demande est complète (formulaire rempli, éventuellement mandat de recherche, description des variables, etc.).

Toute demande d'appariement, quelle que soit sa catégorie, doit contenir les informations suivantes:

- informations sur le mandant (personne, institution)
- description du projet (y c. le but de l'appariement)
- données nécessaires (données à appairer, variables; pour les données externes: métainformations et bases légales)
- résultats visés (éventuellement publication prévue)
- délais
- données supplémentaires (type d'appariement - appariement unique ou récurrent, outils informatiques ou programmes spéciaux, etc.).

Documentation:

Toutes les demandes d'appariement et les appariements réalisés sont documentés. La documentation comprend les éléments suivants: •demande •avis •décision •contrat de protection des données ou règlement de traitement (selon le niveau de protection et la catégorie d'appariement) •

La documentation est gérée dans GEVER..

Appariements systématiques destinés à la production statistique et appariements longitudinaux

L'annexe de l'ordonnance sur les relevés statistiques mentionne les appariements destinés à la production statistique et les appariements longitudinaux. Comme tous les autres appariements, ils doivent être documentés et approuvés par le directeur (voir le point 9.1). Ces appariements sont généralement approuvés au moment de l'adoption de la base conceptuelle, avant qu'ils soient inscrits dans l'annexe de l'ordonnance sur les relevés statistiques et qu'ils fassent l'objet d'un règlement de traitement (s'ils présentent un niveau de protection 3).

Demandes d'appariement répétées

Lorsqu'un appariement a déjà fait l'objet d'une demande similaire, il est possible de renvoyer à la documentation fournie précédemment. Les éventuelles différences feront l'objet d'une description détaillée.

Durée de l'autorisation

Appariements pour la production statistique : l'autorisation et l'accès aux identificateurs ou aux clés nécessaires peuvent être accordés pour une certaine durée et renouvelés périodiquement. La durée souhaitée doit être indiquée dans la demande d'appariement; elle ne peut dépasser cinq ans (voir chap. 10.2.2).

Appariements pour des externes : la durée de l'utilisation des données est définie par contrat. Au besoin, il est possible de demander une prolongation.

9.3 Règlement de traitement

Comme nous l'avons expliqué au chap. 5, l'appariement en tant que tel ne requiert pas un règlement de traitement spécifique, qui viendrait compléter les présentes directives et les dispositions légales applicables. La nécessité d'un tel règlement est déterminée par le niveau de protection dont bénéficient les données d'input et les données d'output. Si ces données présentent un niveau de protection 3, elles font déjà l'objet d'un règlement de protection, destiné notamment à empêcher qu'on puisse y accéder de manière non autorisée. Leur appariement ne requiert donc pas l'établissement d'un règlement de traitement supplémentaire. L'appariement doit cependant être documenté de manière détaillée et transparente. Le formulaire de demande notamment doit indiquer les variables à appairer en particulier et le nom des personnes qui doivent pouvoir accéder aux données dé-pseudonymisées, et combien de temps elles doivent y être autorisées. Pour les données présentant un niveau de protection de 0 à 2, il n'existe en principe pas de règlement de traitement. Dans ces cas, on analysera le besoin de protection pour déterminer si le nouveau jeu de données (appariées) nécessite l'établissement d'un règlement de traitement. Cela peut être le cas si les données d'output doivent être attribuées au niveau de protection 3.

L'annexe 3 contient un modèle de règlement de traitement pour les données de ce niveau, incluant un module pour l'appariement.

9.4 Association de tiers aux travaux d'appariement

Selon l'art. 13k, al. 3, de l'ordonnance sur les relevés statistiques, l'OFS peut associer des tiers (personnes externes à l'OFS) au processus d'appariement pour limiter les coûts et à des fins d'efficacité. La dé-pseudonymisation et l'établissement de l'indicateur spécifique au projet sont toujours réalisés par l'OFS. Si des tiers sont associés à d'autres étapes du processus l'OFS prend les mesures organisationnelles et techniques nécessaires pour garantir la protection et la sécurité des données (voir chap. **Erreur ! Source du renvoi introuvable.**). Dans certains cas, il peut s'avérer pertinent, pour des raisons organisationnelles ou pour assurer la protection des données, de faire effectuer ces travaux dans les locaux de l'OFS à une place de travail sécurisée.

Mesures organisationnelles

- Avant le début des opérations d'appariement associant des tiers, l'OFS signe avec ces derniers un contrat détaillant la procédure, les droits et les obligations des parties et les dispositions régissant la protection des données (contrat de type 2).
- Les responsables de secteur veillent à ce que les indicateurs spécifiques à un projet (voir ch. 10.3.1) et les données nécessaires soient préparés comme il se doit.
- Si des tiers doivent effectuer certains travaux à l'OFS : les responsables de secteur accueillent les externes chargés de l'appariement et les accompagnent à la place de travail prévue à cet effet. Ils règlent les formalités finales et la remise des données appariées (moyennant la signature d'un contrat de protection des données).
- Des spécialistes de l'OFS se tiennent à disposition pour apporter des conseils et un soutien technique.

Mesures techniques

La section IT de l'OFS dispose de places de travail dotées d'un keystore qui sont préparées spécialement pour les travaux d'appariement devant être effectués à l'OFS. Les mesures suivantes sont prises dans de tels cas :

- La place de travail est équipée d'un ordinateur de bureau suffisamment puissant, sur lequel sont installés les logiciels requis. L'ordinateur ne dispose d'aucune connexion ni interface avec l'extérieur. Les tiers associés ne peuvent donc pas exporter des données, ni par un support physique (clé USB, CD, etc.), ni par le réseau (Internet, Intranet, mail, etc.).

Directives sur l'appariement

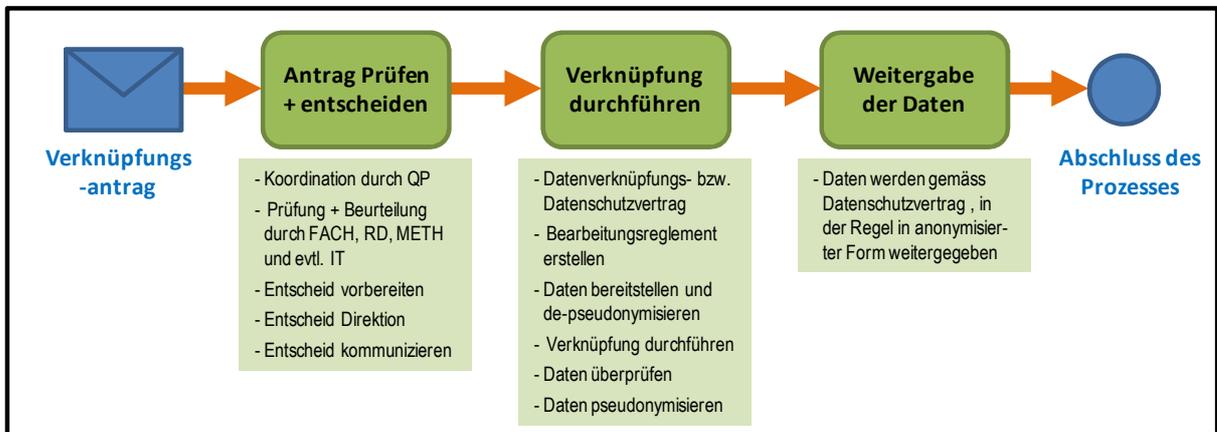
- Les données nécessaires à l'appariement (données d'input) peuvent en tout temps être importées dans l'ordinateur via la clé USB. Les données obtenues (données d'output) ne peuvent être exportées que par l'administrateur-système responsable. Pour l'exportation, il est nécessaire de recourir à l'OSS de l'OFIT. L'utilisation de telles places de travail fermées est réduite au minimum.

10 Processus d'appariement

(voir aussi l'annexe 4 Description du processus)

10.1 Remarques préliminaires

L'appariement suit normalement les étapes suivantes:



Le processus comporte cinq étapes:

- demande d'appariement,
- Examen de la demande, avec décision indiquant si l'appariement peut être effectué, de quelle manière et quand,
- réalisation de l'appariement,
- remise des données appariées,
- clôture du processus d'appariement.

10.2 Examen de la demande et décision

10.2.1 Examen de la demande

L'appréciation d'une demande d'appariement externe est effectuée par les sections de production concernées (SPROD; pour les demandes internes, il suffit d'une demande signée par la direction de la division), par la section Méthodes statistiques (METH) et par le service juridique (RD) (voir aussi l'art. 2, al. 2, de l'ordonnance sur les appariements).

Les rapports d'évaluation sont coordonnés par le service spécialisé en appariements. Ils comprennent deux volets:

- 1) une appréciation du projet fondée sur les spécificités du domaine.
- 2) une demande/recommandation à l'intention de l'instance qui décide.

En premier lieu, la demande est transmise pour information aux chef/fes de division responsables. Puis les sections de production concernées (SPROD) procèdent à l'évaluation du projet. Celle-ci est effectuée parallèlement à l'évaluation méthodologique et juridique du projet par METH et RD. SPROD procède à l'évaluation du projet. Son résultat est transmis avec la demande, par la voie hiérarchique (⇒SC ⇒AC), à METH et à RD (éventuellement à IT) pour une évaluation méthodologique et juridique (et technique dans le cas d'IT). Son résultat est indiqué par écrit dans le formulaire «Demande, appréciation et décision».

Les prises de position de SPROD, METH et RD sont résumées par le *service spécialisé en appariement*, qui les transmet aux chef/fes de division responsables, lesquelles soumettent la demande au

Directives sur l'appariement

conseil de direction. Si la demande n'est pas contestée par un membre du conseil de direction, le document est remis à la direction, qui prend une décision.

Si une demande émane d'un autre office fédéral, on prendra immédiatement contact avec son auteur pour définir une procédure adaptée au projet.

Evaluations

Prise de position SERVICE SPECIALISE EN APPARIEMENT : la demande est-elle complète ? Remarques ?

Evaluation par SPROD: les données et les identificateurs sont-ils disponibles? L'appariement est-il pertinent, a-t-il un sens ? Est-il réalisable (du point de vue méthodologique, technique)? S'agit-il de données sensibles (input/output)? Quel est le coût de l'appariement/quelles sont les ressources disponibles? Procédure (qui fait quoi, quand et comment)? Conditions, réserves/restrictions (dispositions spéciales, utilisation des données de l'output, etc.)?

Evaluation par METH: le projet est-il correct/défendable du point de vue méthodologique? Les résultats ont-ils un sens? Y a-t-il des réserves/restrictions?

Evaluation par RD: les conditions juridiques sont-elles réunies? Evaluation de la sensibilité des données d'input et d'output, réserves, conditions et mesures spéciales.

Dans quelques rares cas, (par ex. enregistrement des données dans un cloud), une évaluation de la part de la section IT est aussi nécessaire.

(r⇒voir l'annexe 5, formulaire

10.2.2 Décision de la direction

Sur la base de la demande/de la recommandation, la direction décide si l'appariement peut avoir lieu, quelle forme il doit prendre, si des allègements sont possible ou s'il faut au contraire définir des réserves (par ex. mesures spéciales, exigences supplémentaires, réutilisation de l'appariement, conservation ou suppression des données, calendrier).

- Les décisions sont prises par écrit.
- Les décisions et les évaluations sont communiquées par écrit au demandeur par la section responsable. Tout refus est motivé.
Si le demandeur estime injustifiés les motifs du refus, il peut demander que la décision soit reconsidérée.
- Une décision positive (écrite) de la direction autorise la *section IT* à générer et à remettre les clés, à la condition que les réserves émises soient respectées.

(⇒voir l'annexe 5 Formulaire

Directives sur l'appariement

Possibles réserves et allègements dans la décision de la direction:

Réserves:

Dans le cas de données sensibles, la décision peut imposer

- des mesures (de sécurité) spéciales lors de la réalisation des appariements,
- des restrictions à la remise ou à l'utilisation des données.

Si les appariements engendrent un coût élevé ou nécessitent un savoir-faire particulier, l'OFS peut exiger qu'ils soient réalisés par le mandant lui-même, sous certaines conditions (voir 10.3.2).

Allègements:

- Lorsqu'un appariement est destiné à la production d'une statistique de routine, une autorisation peut être délivrée pour une certaine durée. Sa durée est déterminée d'après la nature de la production statistique et la sensibilité des données. La décision à cet égard revient à la direction.
- Des allègements peuvent être décidés en fonction de la sensibilité des données, du genre de mandant ou de demande d'appariement. La décision à cet égard revient à la direction.

10.3 Réalisation de l'appariement

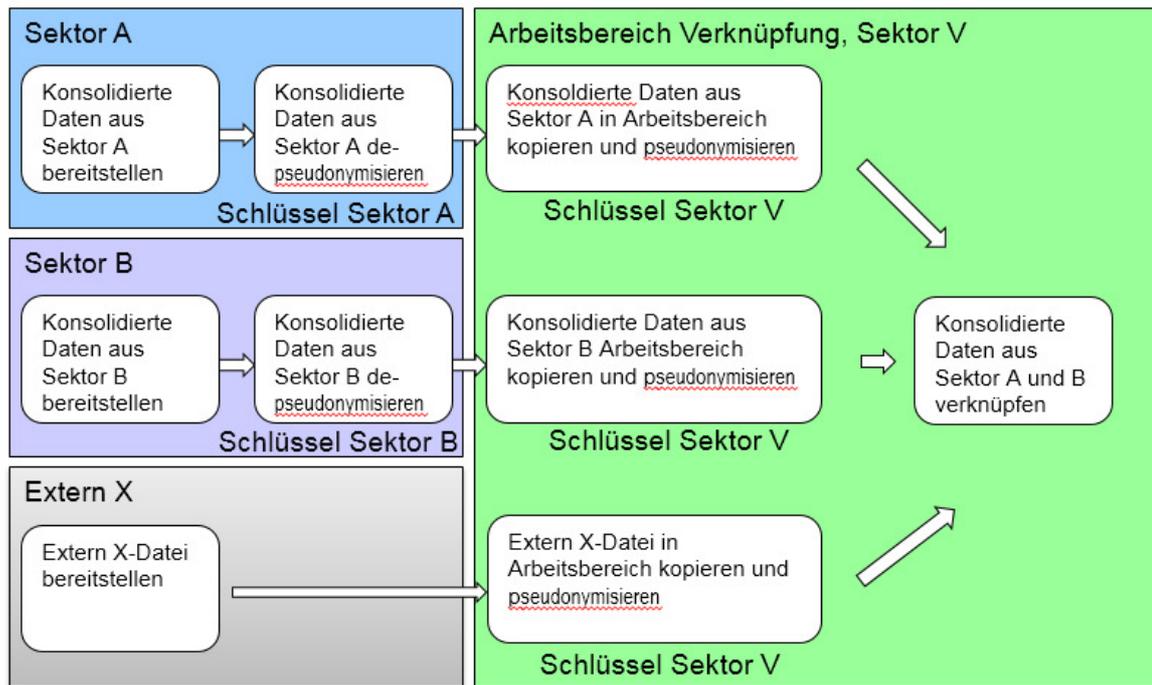
L'appariement est dirigé par la section responsable des données. Lorsque les données ressortissent à différentes sections, ces dernières se répartissent les rôles et les responsabilités.

10.3.1 Appariement par l'OFS

Les appariements sont normalement effectués par l'OFS. La dé-pseudonymisation et l'établissement d'un identificateur spécifique au projet sont toujours réalisées par l'OFS. Idéalement, les appariements comportent les étapes suivantes:

- **Mise à disposition des données:** Les données de l'OFS sont extraites des banques de données et regroupées dans un espace de travail dédié à l'appariement préparé par la section Informatique.
- **Dé-pseudonymisation des données:** Les données pseudonymisées de l'OFS sont dé-pseudonymisées au moyen des clés fournies par le service spécialisé en appariement.
- **Appariement et pseudonymisation des données:** Les données sont pseudonymisées dans l'environnement informatique. Les données individuelles sont combinées grâce aux identificateurs communs spécifiques au projet

Directives sur l'appariement



Ce schéma illustre les domaines de responsabilité des différents secteurs, mais non l'emplacement physique de la banque de données ou du support de stockage. Il est à noter que les données dé-pseudonymisées et les données pseudonymisées ne doivent pas être stockées au même endroit.

10.3.2 Association de tiers au processus d'appariement

L'OFS peut associer des tiers au processus d'appariement à des fins d'économie, d'efficacité ou de partage des connaissances (voir l'art. 13k, al. 3, de l'ordonnance sur les relevés statistiques). La participation de tiers doit être réglée par contrat (voir le contrat de protection de données de type 2).

Mise à disposition des données:

Les collaborateurs responsables à l'OFS dé-pseudonymisent les données de l'OFS nécessaires à l'appariement et les mettent à disposition.

L'OFS réceptionne les données externes (données tierces et autres données) et les collaborateurs responsables les mettent à disposition. Ces données contiennent normalement un identificateur permettant l'appariement.

Ces données (données de l'OFS, données tierces et autres données) sont dotées d'un identificateur approprié pour l'appariement.

Les identificateurs d'origine sont convertis en un nouvel identificateur spécifique au projet. Les données ainsi pseudonymisées sont transmises aux utilisateurs par le biais d'un canal sécurisé.

Appariement concret des données Les tiers appariement les données individuelles au moyen de l'identificateur commun spécifique au projet. Un fichier contenant les données appariées est créé.

Dans des cas justifiés : mettre à disposition une place de travail pour des participants externes (voir chap. 1.1).

Dans certains cas, il est nécessaire, pour des raisons techniques, que des tiers assurent certaines tâches liées à l'appariement dans les locaux de l'OFS. C'est par exemple le cas lorsque des données tierces ou des d'autres données ne sont pas munies d'identificateur approprié pour l'appariement et que l'on applique des méthodes probabilistes ou étapes manuelles demandant beaucoup

Directives sur l'appariement

de temps. On peut citer comme exemple concret la recherche de personnes à l'aide des critères *Prénom, nom et commune de domicile* : ce procédé implique le développement d'une méthode et éventuellement certaines sélections manuelles qui prennent du temps et relèvent de la responsabilité des utilisateurs des données.

Si de tels travaux d'appariement sont accomplis à l'OFS par des personnes externes, les résultats sont encore éventuellement pseudonymisés et dans tous les cas contrôlés par la section de production responsable. Celle-ci examine notamment les points suivants:

- Les travaux d'appariement ont-ils été réalisés conformément au contrat?
- Le fichier qui en résulte correspond-il à l'output convenu par contrat?

L'OFS ne contrôle pas dans les détails la précision du contenu et de la méthodologie des données obtenues. Cette précision relève de la responsabilité du demandeur.

Que certaines étapes de travail soient assurées ou non par des tiers (à l'OFS ou non), les données fournies aux utilisateurs des données au terme du processus d'appariement sont sans exception pseudonymisées comme décrit au point 10.3.1.

10.3.3 Appariement par des services cantonaux ou communaux

Les services statistiques cantonaux ou communaux peuvent être autorisés, à certaines conditions, à réaliser eux-mêmes des appariements avec des données de la statistique fédérale – également sous une forme non pseudonymisée - (voir l'art. 14a, al. 2, LSF, l'art. 13j de l'ordonnance sur les relevés statistiques et l'art. 5 de l'ordonnance du DFI sur l'appariement de données). Si ces conditions sont remplies, qu'un règlement de traitement adéquat a été établi et qu'un contrat de protection des données et d'appariement a été signé par les deux parties, la section de production responsable livre les données éventuellement requises (avec un identificateur approprié) et vérifie que les accords contractuels relatifs à la protection et à l'utilisation des données sont respectés.

10.4 Remise de données appariées

La remise de données appariées est soumise aux mêmes règles que la remise des autres données individuelles. On observera à cet égard, outre les dispositions de la loi et des ordonnances, les «*Directives relatives à la protection des données en cas de transmission de données individuelles à des tiers*» (état au 18.10.2006).

Les détails de la remise de données individuelles sont à régler dans tous les cas par contrat.

10.5 Clôture du projet d'appariement et effacement des données appariées

La communication des données appariées aux fins prévues clôt le processus d'appariement. Il reste à observer les obligations suivantes:

Fin de l'autorisation à utiliser les clés d'appariement remises par la section IT. (voir annexe 5: Formulaire)

Dans l'intérêt de la protection des données, les données appariées doivent également être détruites après la fin de leur exploitation statistique si elles contiennent des *données sensibles ou des profils de la personnalité* (voir art. 13m de l'ordonnance sur les relevés statistiques). L'option de détruire ou d'effacer les données est réglée dans la décision d'appariement (voir chap. 10.2.2), à la rubrique «*Spécifications, réserves*».

Facteurs spéciaux

Différents facteurs peuvent influencer concrètement sur le processus d'appariement et doivent être pris en compte au moment d'évaluer sa pertinence. Il s'agit notamment des suivants:

► **Demandeur.** On distingue les catégories suivantes:

- les demandeurs internes (unités organisationnelles de l'OFS),
- les demandeurs externes (par ex. autres services fédéraux, cantonaux et communaux, instituts de recherche, tiers),
- les offices statistiques des cantons et des communes qui font partie des organes habilités à effectuer des appariements.

► **Données à appairier** (par ex. nature, origine, sensibilité des données). En ce qui concerne l'origine des données, on opère les distinctions suivantes (voir annexe 1 Définitions, Concept SIS):

- A)** données de l'OFS (données dont l'OFS est propriétaire) ⇒ données de l'OFS.
- B)** données relevées dans le cadre de la LSF et de l'ordonnance sur les relevés statistiques qui ne sont pas la propriété de l'OFS (par ex. données d'autres offices fédéraux) ⇒ **données tierces**⁷.
- C)** données qui ne sont pas relevées en vertu de la LSF (par ex. données administratives de services de l'administration ou d'autorités; données relevées par des chercheurs ou d'autres tiers) ⇒ **autres** données.

Ces trois types de données peuvent être appariées dans les combinaisons suivantes:

- **A** avec **A** • **A** avec **B** • **A** avec **C** • **B** avec **B** • **B** avec **C**

L'OFS n'effectue pas d'appariements combinant C avec C, qui ne sont pas soumis à la présente réglementation. De tels appariements nécessitent des bases légales spéciales.

Ces différentes combinaisons rendent possibles une foule de variantes qu'il n'est pas utile de décrire dans les détails. Nous nous contentons de présenter ci-après une variante «standard», en indiquant les ajouts ou variations possibles.

⁷ Dans le concept SIS, il n'est question que d'autres données. L'annexe de l'ordonnance sur les relevés statistiques précise ce qu'il faut entendre par données tierces.

11 Annexes

11.1 Annexe 1: Définitions (concept SIS)

Définitions (concepts SIS, pp. 48 à 51 /148)

Les domaines suivants s'inscrivent dans la chaîne de valorisation et correspondent aux concepts du SIS:

Données d'input

Par données d'input (aussi appelées données brutes), on entend les données qui proviennent des fournisseurs de données et de l'exécution des relevés de l'OFS et qui ont été acceptées par l'OFS. Leur principale caractéristique est qu'elles n'ont pas encore été traitées et qu'elles se présentent donc sous une forme brute. Elles sont enregistrées sous cette forme à des fins de traçabilité.

A l'OFS, on fait normalement une distinction entre les données d'input primaires (qui sont destinées à une utilisation statistique) et les données d'input secondaires (collectées au départ à d'autres fins mais pouvant aussi être utilisées à des fins statistiques).

Nous utilisons toutefois un autre critère pour distinguer les données d'input, celui du responsable des données. Selon ce critère, on distingue les deux catégories suivantes:

- a) Données résultant de relevés exécutés ou mandatés par l'OFS:
Il s'agit de données relevées par l'OFS lui-même ou par des institutions externes mandatées par l'OFS. L'OFS dispose de ces données sous une forme brute et doit encore les contrôler et les valider.
Ces données peuvent être des données primaires (données issues des relevés réalisés ou mandatés par l'OFS à des fins statistiques) ou des données secondaires (en particulier données des registres tenus par l'OFS).
- b) Données externes et données de relevé (autres données)
Ces données ne proviennent pas de relevés effectués par l'OFS ou sur mandat de ce dernier, mais par d'autres offices, institutions publiques ou autres, qui livrent des données déjà traitées (généralement validées, plus rarement agrégées).
Ces données n'ayant pas été collectées à des fins statistiques, elles sont souvent incomplètes. Il incombe à l'OFS d'en combler les lacunes. Les données externes déjà traitées (en particulier celles qui sont agrégées) ne peuvent en revanche plus être validées.
Parmi ces "autres données", on trouve aussi bien les données primaires qui n'ont pas été saisies par l'OFS ou ses mandataires et toutes les données secondaires (qui ne visent pas des fins statistiques). Les données d'input comprennent plus largement les données concernant les produits des fournisseurs de données externes.

Etant donné que ces données (modifiées pour servir des fins statistiques) n'appartiennent plus à leur fournisseur primaire, l'OFS en devient leur propriétaire dans les limites d'une utilisation statistique.

Données consolidées

Les données consolidées sont des données individuelles ou microdonnées.

En font partie les microdonnées préparées et pseudonymisées, le cas échéant complétées de données calculées ou dérivées (par ex. l'âge tiré de l'année de naissance). Les données consolidées comprennent aussi les états intermédiaires ($A_0 \dots A_n$, F) selon la définition du processus de préparation statistique des données (PPSD).

Les *données agrégées* en sont toutefois exclues.

Les données consolidées comprennent les microdonnées préparées individuellement, qui sont protégées par des mesures spéciales.

Directives sur l'appariement

Format physique:

Les données consolidées sont stockées dans le pool de données, pour chaque relevé, sous la forme d'un tableau de données relationnel ordinaire ou d'un tableau multidimensionnel. La consolidation permet d'intégrer d'autres formats dans ce type de présentation.

On peut de la sorte représenter correctement l'ensemble des formats d'input connus, même si cela réduit le nombre de variantes de formats utilisables pour la suite (et donc le nombre de transformations possibles pour construire des compteurs de données destinés à l'analyse). Une autre raison technique de cette démarche est le fait que le pool de données (en tant que banque de données purement relationnelle) n'est utilisable qu'avec Oracle.

Remarque: A ce stade, les données consolidées présentent normalement déjà des propriétés multidimensionnelles. L'implémentation vise à déterminer s'il faut privilégier une représentation (normalisée) sous forme de schéma en étoiles ou en flocons ou un fichier plat pour l'extraction et la transformation des données en cubes ou en compteurs de données. Il ne faut en aucun cas utiliser une représentation propre au producteur des données qui, comme le montre l'expérience, posera problème lors de la conversion.

Données de production

Les données de production sont normalement le résultat d'une analyse ou interprétation effectuée (parfois automatiquement) ; elles servent de base aux étapes suivantes de publication. Elles se composent de données (partiellement) agrégées.

Les données de production comprennent aussi les données agrégées préparées spécialement à l'intention d'un mandant.

Elles se présentent sous forme de tableaux factuels (contenant des renvois vers les dimensions concernées, à savoir les nomenclatures, plages de valeurs, etc.). Tandis que les POE utilisent toute sorte d'outils d'implémentation des cubes et des compteurs de données dans les espaces de travail, les cubes sont stockés dans les espaces de données dans un format déterminé.

Format physique:

- a) **Compteurs de données:** Le format de représentation dans le pool de données est défini comme un schéma relationnel en étoiles (ou en flocons). Lors du transfert à partir du domaine de données vers l'espace de travail, et inversement, ce format permet de transformer les données de manière ouverte et générique à partir de ou vers n'importe quel format de cube.
- b) **Les tableaux** sont stockés sous la forme de tableaux relationnels (ils font partie intégrante de la banque de données).

Données destinées à la publication

Ce domaine de données contient trois sortes de données:

1. Toutes les données standards dont l'utilisation publique est autorisée. Ces données doivent satisfaire aux exigences de la protection des données.

Les données destinées à la publication traversent un processus de production sur le fond avant d'être soumises à un traitement rédactionnel pour donner naissance à un «produit de visualisation». Elles peuvent aussi déboucher sur des données consolidées, des données de produit ou des métadonnées, éventuellement accompagnées par un texte explicatif (registres, nomenclatures, etc.).

2. Toutes les données fournies directement à des tiers (articles individualisés). Il peut s'agir, à l'instar du point 1, de données consolidées, de données de produit ou de métadonnées, mais elles diffèrent du fait qu'elles ne sont pas publiques, mais uniquement accessibles aux partenaires impliqués.
3. Ce domaine contient également les «objets d'information», même s'il s'agit de produits à considérer comme «semi-finis».

11.2 Annexe 2: Bases légales

1. Loi sur la statistique fédérale (LSF; SR 431.01)

Art. 14a1 Appariement de données

1 Pour exécuter ses tâches en matière de statistiques, l'office peut appairer des données à condition de les rendre anonymes. Si des données sensibles sont appariées ou si l'appariement de données permet d'établir des profils de la personnalité, les données appariées doivent être effacées une fois les travaux statistiques d'exploitation terminés. Le Conseil fédéral règle les modalités.

2 Les services cantonaux et communaux de statistique ne sont autorisés à appairer les données de l'office avec d'autres données pour exécuter leurs tâches en matière de statistiques qu'avec l'accord écrit de ce dernier et aux conditions qu'il aura fixées.

2. Loi sur l'harmonisation de registres (LHR; SR 431.02)

Art. 16 Utilisation des données par l'office à des fins statistiques, de recherche et de planification

1 L'office peut utiliser les données pour effectuer des relevés ou des exploitations statistiques.

2 Il peut tirer des échantillons à partir des données pour effectuer des relevés statistiques.

3 Il peut utiliser les données mentionnées à l'art. 6, let. a à h, j, k et m, pour constituer un répertoire d'adresses pour l'exécution de relevés statistiques.

4 Il peut, pour remplir ses tâches en matière de statistiques, appairer les données dépourvues de désignation de personne avec celles du RegBL et du Registre des entreprises et des établissements (REE) et les conserver durablement.

3. loi fédérale sur la protection des données (LPD; SR 235.1)

Art. 22 Traitement de données sur mandat

1 Les organes fédéraux sont en droit de traiter des données personnelles à des fins ne se rapportant pas à des personnes, notamment dans le cadre de la recherche, de la planification ou de la statistique, aux conditions suivantes:

- a. les données sont rendues anonymes dès que le but du traitement le permet;
- b. le destinataire ne communique les données à des tiers qu'avec le consentement de l'organe fédéral qui les lui a transmises;
- c. les résultats du traitement sont publiés sous une forme ne permettant pas d'identifier les personnes concernées.

2 Les dispositions suivantes ne sont pas applicables en la matière:

- a. art. 4, al. 3, relatif au but du traitement;
- b. art. 17, al. 2, relatif à la base juridique pour le traitement de données sensibles et de profils de la personnalité; et
- c. art. 19, al. 1, relatif à la communication de données personnelles.

4. Ordonnance concernant l'exécution des relevés statistiques fédéraux (Ordonnance sur les relevés statistiques); SR 431.012.1)

Section 2a Appariements de données

Art. 13h Définition

On entend par appariement de données le fait de relier des données provenant de sources différentes, telles que données d'enquêtes, données de registres, données administratives ou données de mesures.

Art. 13i Principes

1 Les appariements de données servent à obtenir des informations statistiques sans avoir à effectuer de relevés.

2 Ils ne sont réalisés que s'ils sont appropriés et nécessaires à l'accomplissement de travaux statistiques.

Art. 13j Conditions

1 Seules sont appariées des données qui présentent les caractéristiques et la qualité requises pour un traitement statistique.

2 Pour exécuter ses tâches statistiques, l'OFS peut appairer aussi bien ses propres données que des données sur lesquelles il n'a aucun droit (données tierces).

3 Le fournisseur de données tierces qui donne à l'OFS le mandat d'appairer ces données (art. 13k) doit fournir la preuve que:

- a. le relevé des données, leur transmission à l'OFS et leur appariement sont conformes à la loi, et que
- b. ces données présentent la qualité statistique requise.

4 Les services statistiques des cantons et des communes sont autorisés, pour exécuter leurs tâches statistiques, à appairer des données statistiques de l'OFS entre elles ou avec leurs propres données, pour autant qu'ils s'engagent, dans un contrat de protection des données:

- a. à garantir la protection des données au même degré que l'OFS;
- b. à ne pas communiquer les données de l'OFS sans son accord écrit;
- c. à garantir leur indépendance vis-à-vis des organes exécutifs;
- d. à édicter et à mettre en oeuvre un règlement de traitement des données;
- e. à prendre des mesures suffisantes pour garantir la sécurité et la protection des données;

Directives sur l'appariement

f. à respecter les normes de bonnes pratiques de la statistique.

Art. 13k Appariement de données sur mandat de tiers

- 1 L'OFS effectue des appariements sur mandat de tiers à des fins ne se rapportant pas à des personnes, notamment à des fins de recherche, de planification et de statistique, dans le cadre d'un contrat de protection des données, selon ses moyens techniques, organisationnels et humains. Il appuie en particulier les projets d'appariement de la Confédération et des cantons.
- 2 L'indemnisation est régie par l'ordonnance du 25 juin 2003 sur les émoluments et indemnités perçus pour les prestations de services statistiques des unités administratives de la Confédération¹.
- 3 Pour limiter les coûts et à des fins d'efficacité, l'OFS peut, pour certaines tâches, associer le mandant au processus d'appariement. Ces tâches sont clairement définies dans un contrat de protection des données.

Art. 13l Communication de données appariées

Lorsque la loi prévoit que les données peuvent être communiquées à des fins ne se rapportant pas à des personnes, notamment à des fins de recherche, de planification et de statistique, l'OFS peut, aux conditions fixées à l'art. 9, communiquer des données appariées à des services de statistique ou de recherche de la Confédération ou à des tiers.

Art. 13m Destruction des données appariées

- 1 Les données appariées sont détruites après la fin de leur exploitation statistique si elles contiennent des données sensibles ou des profils de la personnalité.
- 2 Les autres données appariées peuvent être réutilisées pour des travaux statistiques ultérieurs.

Art. 13n Indication des appariements de données

Les statistiques qui donnent lieu à des appariements de données systématiques sont spécifiées en annexe.

5. Ordonnance du DFI concernant l'appariement de données statistiques (Ordonnance sur l'appariement de données); SR 431.012.13)

du 17 décembre 2013 (Etat le 15 janvier 2014)

Le Département fédéral de l'intérieur (DFI),

vu l'art. 14 de l'ordonnance du 30 juin 1993 sur les relevés statistiques¹,

arrête:

Art. 1 Objet

- 1 La présente ordonnance règle l'organisation et le déroulement de l'appariement de données statistiques par l'Office fédéral de la statistique (OFS), la protection et la sécurité des données, ainsi que la participation de tiers au processus d'appariement et les conditions de cette participation.
- 2 Elle règle les conditions que doivent remplir les services statistiques des cantons et des communes impliqués.

Art. 2 Organisation et déroulement des appariements

- 1 Avant toute réalisation d'un appariement à l'OFS, une demande écrite et motivée est adressée à la direction de l'OFS.
- 2 Les sections concernées, le service de méthodologie et le service juridique de l'OFS examinent la demande d'appariement en vérifiant si elle est valable et réalisable des points de vue technique, méthodologique et juridique.
- 3 La compétence de décider de l'opportunité d'un appariement et de l'autoriser appartient à la direction de l'OFS.
- 4 L'état-major de direction est responsable de la gestion et de la remise des clés de liaison (key management) nécessaires à la réalisation et à la surveillance des appariements.

Art. 3 Protection et sécurité des données

- 1 Les clés de liaison nécessaires à la réalisation des appariements sont conservées de manière centralisée et particulièrement sécurisée.
- 2 L'utilisation d'une clé de liaison est soumise à une autorisation écrite de la direction de l'OFS.
- 3 L'état-major de direction remet les clés de liaison aux personnes habilitées à réaliser les appariements. Cette remise est convenue dans un procès-verbal.
- 4 L'OFS s'assure que les appariements sont réalisés en utilisant les moyens techniques les plus récents et en respectant les bonnes pratiques statistiques.

Art. 4 Participation de tiers au processus de l'appariement

- 1 La forme de la participation de tiers à la réalisation des appariements et le rôle joué par ces derniers sont réglés au préalable dans un contrat de protection des données, de même que l'utilisation des données appariées.
- 2 Les tiers participant à l'appariement exécutent leur travail à une place de travail sécurisée de l'OFS, où il n'est pas possible d'importer ni d'exporter des données.
- 3 L'OFS remet à ces tiers les données appariées après en avoir vérifié la forme et le contenu.

Art. 5 Exigences posées aux services statistiques cantonaux et communaux impliqués

Les services statistiques des cantons et des communes peuvent appier des données de l'OFS aux conditions suivantes:

- a. ils disposent des compétences en matière statistique nécessaires pour appier des données correctement des points de vue du contenu et de la méthode et dans la qualité requise;

Directives sur l'appariement

- b. ils exercent exclusivement une activité statistique, indépendante de tâches de surveillance, d'exécution ou de régulation;
- c. ils garantissent le respect du secret statistique et la protection des données personnelles, et
- d. ils accomplissent leur travail statistique dans le respect des principes de l'indépendance scientifique et de l'objectivité.

Art. 6 Règlement de traitement

L'OFS édicte un règlement de traitement qui règle les autres détails de la réalisation des appariements.

Art. 7 Entrée en vigueur

La présente ordonnance entre en vigueur le 15 janvier 2014.

6. Ordonnance relative à la loi fédérale sur la protection des données (OLPD; SR 235.11)

Art. 21 Règlement de traitement

1 Les organes fédéraux responsables établissent un règlement de traitement pour les fichiers automatisés qui répondent à l'un des critères suivants:

- a. contenir des données personnelles sensibles ou des profils de la personnalité;
- b. être utilisés par plusieurs organes fédéraux;
- c. être accessibles aux cantons, à des autorités étrangères, à des organisations internationales ou à des personnes privées;
- d. être connectés à d'autres fichiers.

2 L'organe fédéral responsable précise dans le règlement de traitement son organisation interne. Il y décrit en particulier les procédures de traitement et de contrôle des données et y intègre les documents relatifs à la planification, à l'élaboration et à la gestion du fichier. Le règlement contient les informations nécessaires à la déclaration des fichiers (art. 16) et les indications suivantes:

- a. l'organe responsable de la protection et de la sécurité des données;
- b. la provenance des données;
- c. les buts dans lesquels des données sont régulièrement communiquées;
- d. les procédures de contrôle et en particulier les mesures techniques et organisationnelles visées à l'art. 20 de la présente ordonnance;
- e. la description des champs de données et des unités d'organisation qui y ont accès;
- f. l'accès des utilisateurs au fichier, ainsi que la nature et l'étendue de cet accès;
- g. les procédures de traitement des données, notamment les procédures de rectification, de blocage, d'anonymisation, de sauvegarde, de conservation, d'archivage ou de destruction des données;
- h. la configuration des moyens informatiques;
- i. la procédure d'exercice du droit d'accès.

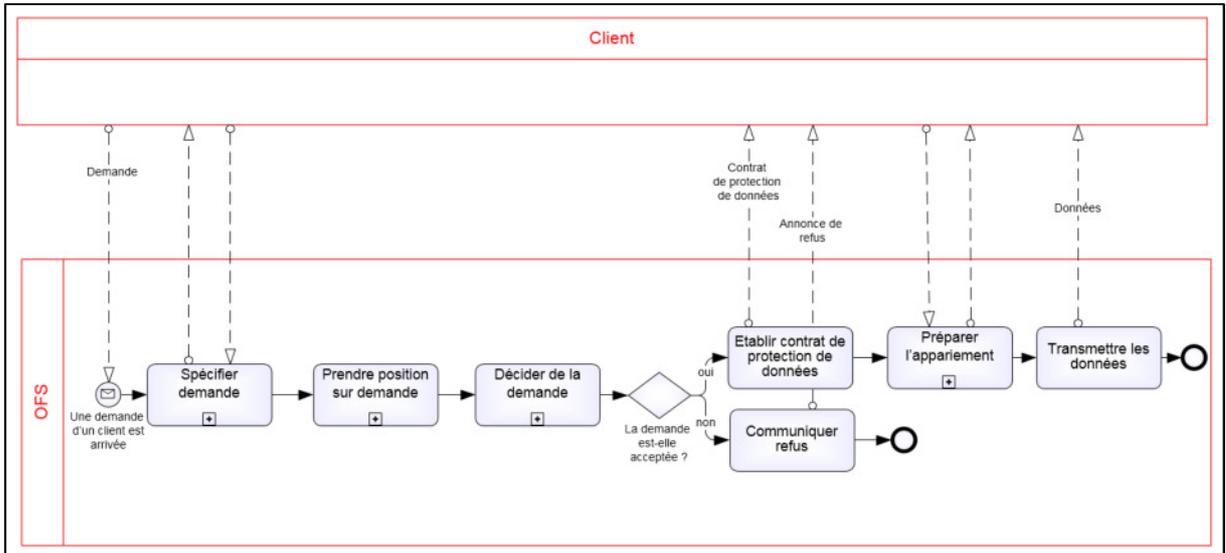
3 Le règlement est régulièrement mis à jour. Il est mis à la disposition des organes chargés du contrôle sous une forme qui leur est intelligible.

11.3 Annexe 3: Règlement de traitement (modèle)

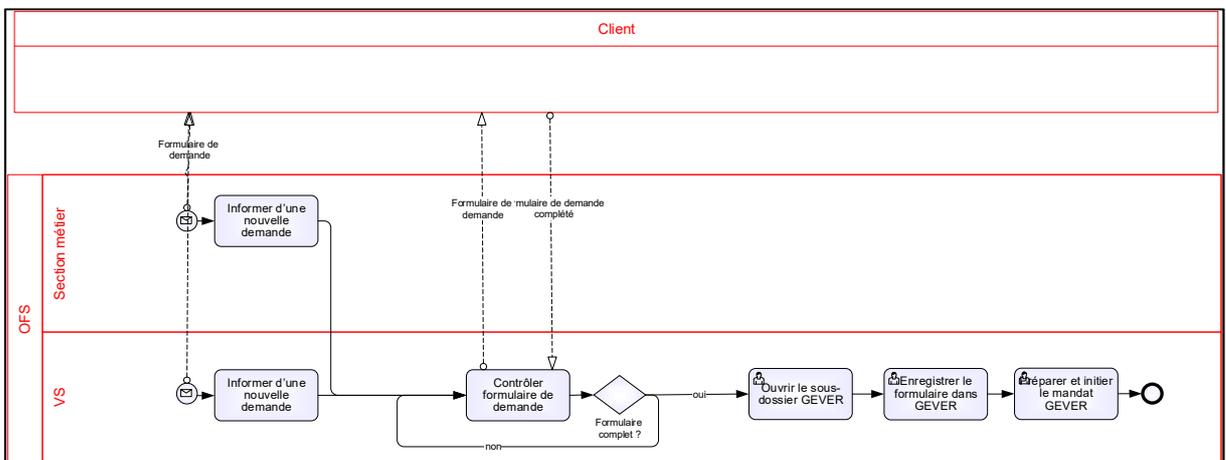
Le modèle est enregistré dans l'archive électronique de l'OFS (uniquement interne à l'OFS) :
<https://gever.edi.intra.admin.ch/edi/mx/COO.2080.104.5.652252?>

11.4 Annexe 4: Description des processus (voir aussi chap. 10)

11.4.1 Traiter une demande d'appariement (vue d'ensemble)

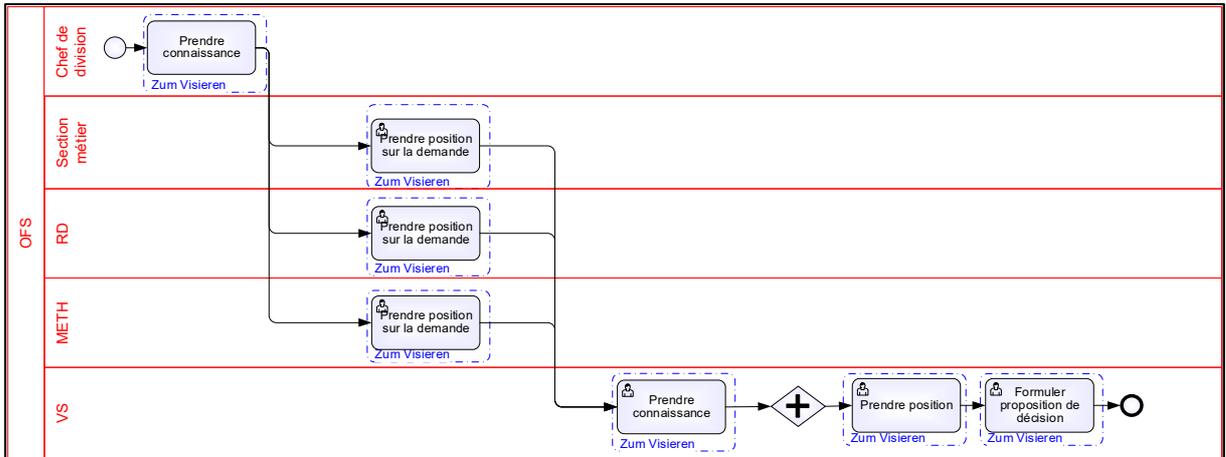


1. Spécifier la demande

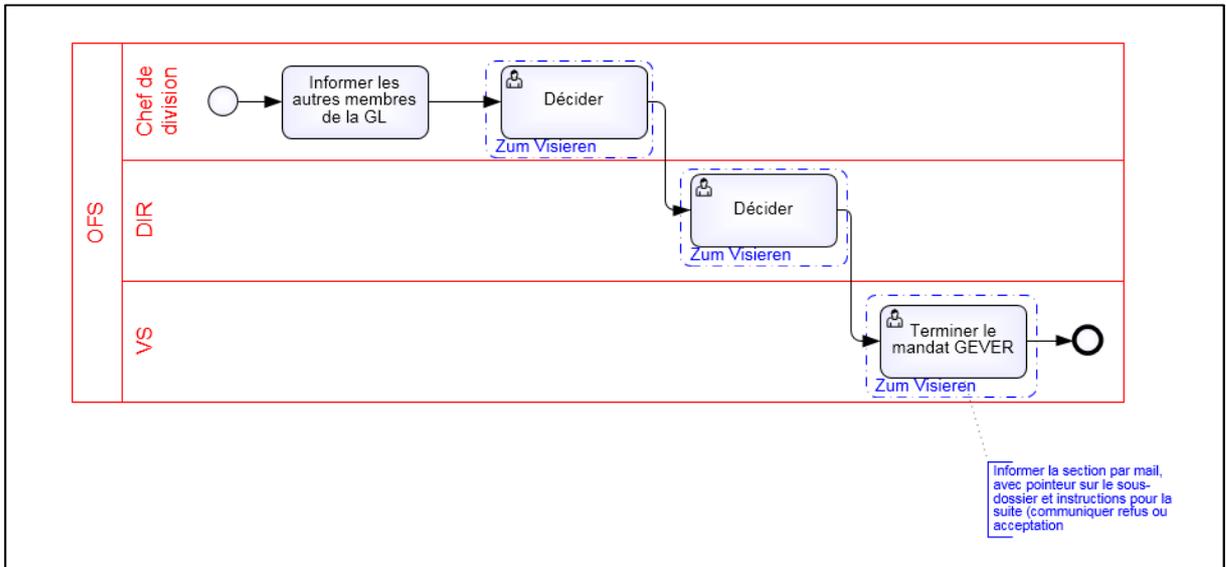


Directives sur l'appariement

2. Prendre position sur la demande

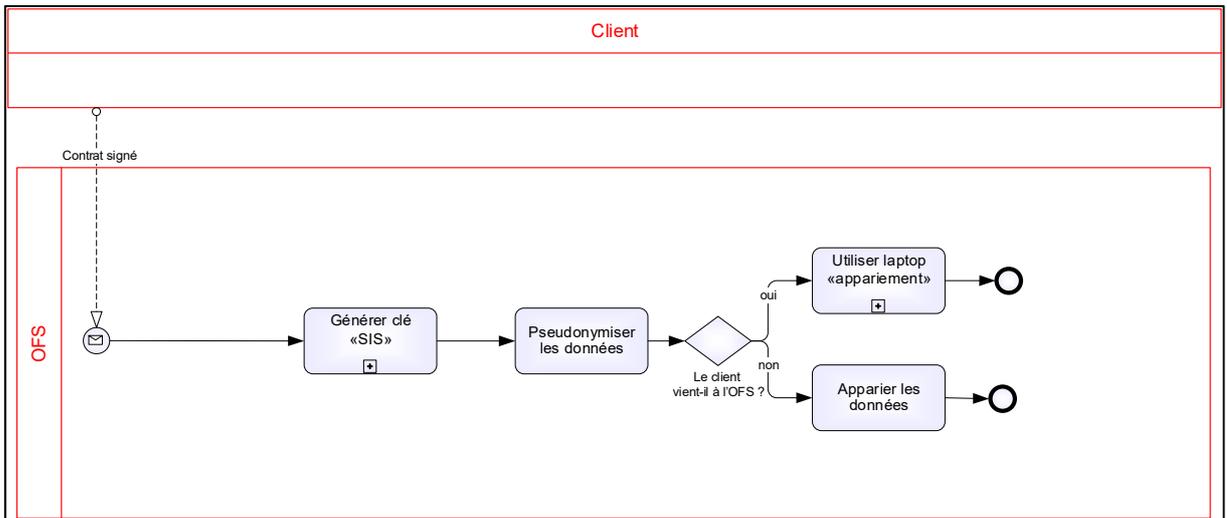


3. Décider de la demande

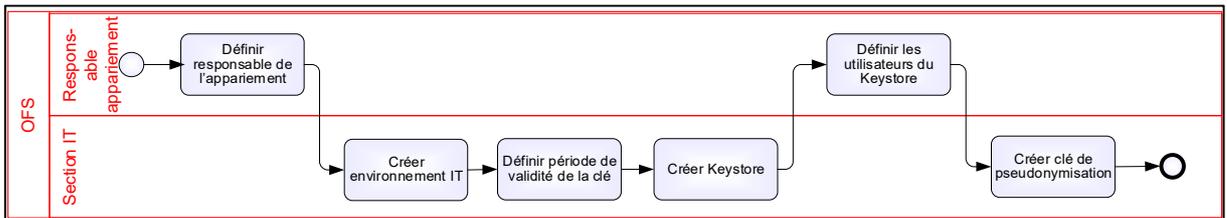


Directives sur l'appariement

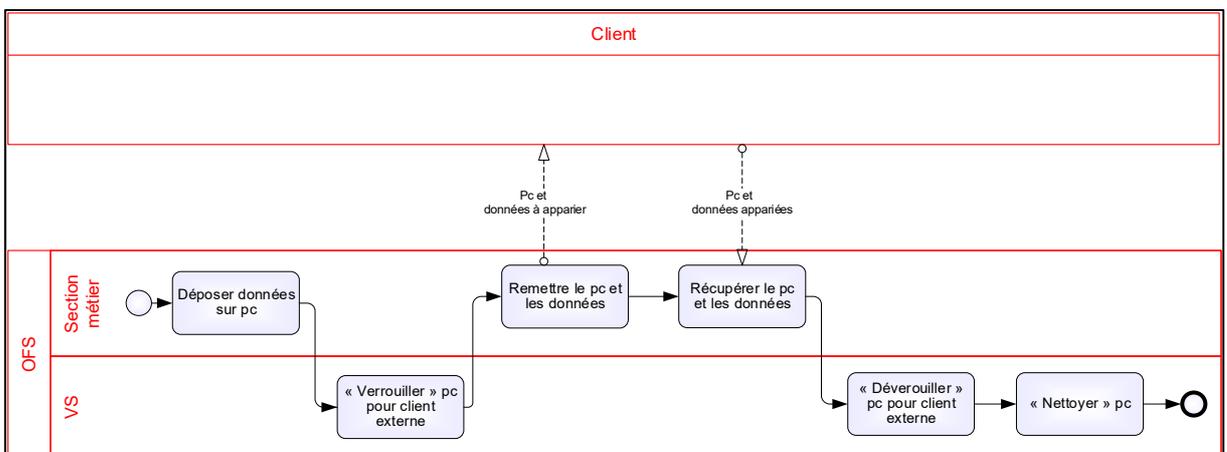
4. Préparer l'appariement



4.1 Générer une clé de pseudonymisation „SIS



4.2 Utiliser un pc « appariement »



VS coordonne le verrouillage/déverrouillage qui est effectué par l'OSS du BIT.

Directives sur l'appariement

Remarque : cette documentation des processus a été réalisée avec MS Visio. Le fichier original peut être consulté ici (uniquement interne à l'OFS) :  [Processus \(214.0-00667/00001\)](#)

11.5 Annexe 5 : Formulaire

Demande, évaluation, décision

<https://www.bfs.admin.ch/bfs/fr/home/services/appariement-donnees/pour-compte-tiers.as-setdetail.7806088.html>

11.6 Annexe 6 : Conservation des données dans des clouds

Les données fournies dans le cadre d'un projet d'appariement ne peuvent pas être conservées dans un cloud par des tiers. Cette interdiction a été décidée lors d'un échange avec le PFPDT et consignée dans un note (uniquement interne à l'OFS) :  [Datenspeicherung in Cloud \(053.0-1/00008\)](#)

Un catalogue de critères, permettant de décider sur d'éventuelles exceptions, a été élaboré (uniquement interne à l'OFS) :  [SpeicherungGeteilteUmgebung_F_final](#)