



Paris, le 2 mars 2009 - N° 73/D130

**Compte rendu de la troisième réunion
du groupe de travail sur les conséquences du profilage des groupes
en matière de secret statistique
du 7 janvier 2009**

(Validé le 21 janvier 2009 lors de la 4ème réunion du groupe)

Présents :

Yvonick	RENARD	MEDEF - Président
François	VINCENT	CFE-CGC - Vice-président
Jacques	FÉRET	Rapporteur
Emmanuelle	FLAMENT-MASCARET	AFEP
Jean-Marie	ROUX	CGT
Jean-Marc	BÉGUIN	DSE INSEE
Gérard	LANG	DCSRI - Suppléant
Jean-Pierre	LE GLÉAU	DCSRI - Titulaire
Jean	LIENHARDT	DSE INSEE
Henri	MARIOTTE	SOeS
Yves	ROBIN	DGCIS
Sébastien	ROUX	CREST

Absents excusés :

Francis	DESMARCHELIER	AFEP
Karine	MERLE	MEDEF
Jean-Paul	VALUET	ANSA

I. PROCHAINES RÉUNIONS

Une prochaine réunion est programmée le 21 janvier 2009 à 14h.
Le lieu en sera précisé par le Cnis.

Madame Emmanuelle Flament-Mascaret participe à ce groupe de travail pour bien manifester l'intérêt et l'importance que l'AFEP attache à ces sujets, M. Desmarchelier étant empêché.

Jean-Marc Béguin, directeur des statistiques d'entreprises de l'Insee, a rejoint ce groupe. Il souhaite apporter des précisions sur la position de l'Insee en matière de profilage à la suite de l'examen du compte rendu de la réunion du 11/12/2008 où deux types de profilages avaient été cités et leur champ respectif examiné.

Il rappelle qu'en droit ce seront toujours les unités légales qui seront redevables d'une réponse aux enquêtes statistiques. L'Insee souhaite bien sûr que la majorité des groupes sollicités contractent un accord amiable, de préférence écrit, avec la statistique publique pour remplacer les filiales par des entreprises profilées pour la réponse aux enquêtes d'entreprise. Il n'est pas sûr que cet accord amiable puisse ou doive prendre dans tous les cas la forme d'une convention. Il ne pense donc pas qu'il faille en faire une condition sine qua non de la réussite d'un profilage. L'expérience des trois premiers profilages montre que cet accord peut prendre la forme de simples échanges de courriers validant les décisions prises en réunion de travail, voire d'un simple accord oral alors que la signature formelle d'une convention risque d'effrayer certains groupes qui feront intervenir leur service juridique. Le travail du groupe en direction de la rédaction d'une convention type est donc plus à entendre comme une check list des points qui devront être traités avec le groupe que comme une convention type stricto sensu. Le rapporteur précise que tel était bien l'esprit dans lequel avait travaillé le groupe.

Par ailleurs, actuellement l'Insee n'a pris des engagements, notamment dans les instances du Cnis, que pour réaliser une centaine de profilages contractualisés à la suite de rencontres bilatérales (cible 1). Il s'agit d'une position prudente compte tenu des moyens que l'Insee peut affecter à ce travail dans une période marquée par des réductions d'effectifs. J.M. Béguin convient qu'il serait plus cohérent avec le récent décret qui définit environ 200 « grandes entreprises » de réaliser plutôt 200 profilages de ce type. Le Service Statistique Public (SSP) fera au mieux avec les moyens dont il disposera mais les engagements du directeur général de l'Insee portent actuellement plutôt sur une centaine.

Pour la quasi totalité des groupes de moins de 500 personnes (environ 38 000 groupes qui forment la cible 2) le profilage sera automatisé et il pense actuellement que dans la très grande majorité des cas, il s'agira d'une procédure additive qui ne pose pas de problème de consolidation. Pour environ 2 000 groupes de taille intermédiaire (cible 3), on peut envisager de faire du semi automatique en utilisant l'information disponible et en introduisant des questions complémentaires dans les enquêtes pour avoir des clés de consolidation.

Si un groupe sollicité dans la cible 1 refuse d'entrer dans cette politique de profilage, le droit commun s'applique. Les filiales continueront d'être enquêtées comme aujourd'hui et le service statistique réalisera lui-même le profilage en utilisant toute l'information disponible (pour les grands groupes on en trouve par exemple beaucoup sur internet) et en sollicitant le groupe pour une validation ex-post sur la plus ou moins grande vraisemblance de la consolidation obtenue. Cette procédure est formellement identique à ce que font tous les statisticiens face au problème de la non réponse.

L'hétérogénéité des statistiques d'entreprise ne disparaîtra pas avec le profilage, mais sera diminuée par rapport à la situation présente. Le profilage vise à améliorer la pertinence des statistiques d'entreprise.

En réponse à une question sur les groupes bancaires, J-M. BEGUIN précise que ces groupes ne sont pas pour le moment dans l'objectif prioritaire. L'Insee est dans l'attente des réflexions du groupe de travail Cnis sur les groupes financiers présidé par M. De Margerie.

Ces précisions apportées figureront au compte rendu de cette réunion et le compte rendu de la réunion du 11/12/2008 est validé.

La discussion est ensuite ouverte en s'appuyant sur la note préparatoire 2008/06 qui ordonnait un certain nombre de questions sur l'application du secret statistique à ces entreprises profilées selon la jurisprudence actuellement appliquée. Cette jurisprudence porte sur la diffusion des données agrégées, la diffusion de données individuelles, les habilitations à l'ensemble des informations soit permanentes, soit limitées.

Les deux règles appliquées actuellement pour interdire **la diffusion d'une donnée agrégée** :

- règle des trois unités , un agrégat n'est diffusable que s'il concerne au moins 3 entreprises,
- règle des 85 % , un agrégat n'est diffusable que si aucune entreprise ne représente 85% ou plus de cet agrégat,

peuvent continuer à s'appliquer en considérant que le terme « entreprise » désigne les unités statistiques observées, c'est à dire les entreprises issues du profilage qui peuvent être soit des entreprises profilées, regroupements de filiales des groupes, soit des unités légales indépendantes n'appartenant pas à un groupe.

A l'occasion de cette discussion, il est précisé que le profilage d'un groupe sera complet : toutes ses filiales sont considérées et chacune est classée dans une *entreprise profilée*¹. Toutefois il se peut qu'une filiale à elle seule soit considérée comme une entreprise profilée si elle jouit d'une autonomie de décision suffisante. Il est même vraisemblable que ce ne sera pas un cas exceptionnel si on se réfère aux 3 premiers profilages réalisés². Quel que soit l'enregistrement de cette situation dans le répertoire statistique, conceptuellement cette entreprise est à classer parmi les entreprises redéfinies.

Les participants notent que la conséquence du profilage sera de réduire la diffusion de résultats agrégés en augmentant le nombre de cas couverts par le secret statistique. Des évaluations³ reposant sur des hypothèses plutôt pessimistes en attendant la réalisation des profilages conduisent à considérer que seuls 4 secteurs de niveaux 700 contiendraient un nombre trop faible d'entreprises pour pouvoir faire l'objet d'une diffusion. Par contre, si l'on veut diffuser des résultats par catégorie de taille d'entreprise, les contraintes sont plus fortes et le groupe de travail du Cnis qui a remis son rapport en novembre 2008 sur ce sujet a considéré qu'il fallait limiter la diffusion de données par catégorie de taille d'entreprises au niveau des divisions de la NAF ; à ce niveau, 9 secteurs⁴ poseront des problèmes de secret statistique sur les grandes entreprises (cf. tableau en annexe).

Bien que l'accroissement du nombre de secrets ne soit pas tel qu'il rendrait ces statistiques inutilisables, il paraît souhaitable de chercher à lever ces secrets par un accord explicite des entreprises concernées comme le prévoit la jurisprudence actuelle. Une telle demande peut en effet aboutir car les grands groupes diffusent souvent des informations nombreuses dans leur communication institutionnelle, soit de leur propre initiative, soit en application des obligations légales sur l'information réglementée définie par l'Autorité des Marchés financiers (AMF). Cependant, cette démarche ne résout pas tous les problèmes puisqu'il y a également des résultats non diffusables portant sur des entreprises de taille intermédiaire voire des PME.

Le GT a ensuite débattu à la fois de la **diffusion de données individuelles d'entreprise et de la communication des contours des entreprises profilées aux utilisateurs de statistiques**. Les deux sujets sont en effet connexes car, si on admet de diffuser des listes d'entreprises dont une partie, notamment les plus grandes, sont des entreprises profilées, la question est alors posée de savoir ce que l'on entend par telle entreprise identifiée par sa seule dénomination par exemple ACCOR HOTELLERIE France, RENAULT DIF ou PSA DAF.

En effet actuellement quand la statistique publique diffuse soit la liste des principales entreprises d'un secteur d'entreprise ou d'une branche d'activité dans l'EAE, soit des listes d'entreprises issues de Sirène, leur identification est celle de l'unité légale dont l'immatriculation est obligatoire dans Sirène et donc bien documentée. Pour les trois premiers profilages réalisés, comme rappelé dans la note de

¹ On utilise faute de mieux ce terme pour désigner les « entreprises » résultant du profilage d'un groupe (littéralement c'est le groupe qui est profilé). Les discussions menées dans le cadre du « groupe Salustro » n'avaient pas permis de trouver un consensus sur ce point de vocabulaire. On parle aussi « d'entreprises redéfinies » ou « reconstituées ».

² Ce serait le cas de Gefco chez Peugeot par exemple.

³ Menées dans le cadre d'un groupe du CNIS qui a travaillé à la préparation du décret 2008-1354 associé à l'article 51 de la LME qui institue 4 catégories de taille d'entreprises. Ces catégories de taille combinent 3 critères : les effectifs, le CA et le total du bilan.

⁴ Ces évaluations ont été réalisées en comptabilisant les entreprises indépendantes et les groupes et non les entreprises indépendantes ou profilées. Il s'agit donc d'une estimation majorée du nombre de résultats confidentiels. Par ailleurs les simulations ont été faites en NAF Rev 1 et sur des classes de taille et non pas des catégories au sens du décret 2008-1354. Mais il s'agit de résultats très structurels et robustes qui ne devraient pas être qualitativement différents en entreprises indépendantes ou profilées, par catégorie de taille et en Naf Rev 2.

travail 2008/02 diffusée pour la 1^{ère} réunion du GT, l'Insee et le Sessi ne publient pas le contour des entreprises profilées mais des indications telles que le nombre de filiales regroupées qui permettent à l'utilisateur de se faire une idée relativement précise de ce que signifie RENAULT Division Industrielle France ou PSA Division Automobile France. L'utilisateur avisé peut par exemple constater que REAGROUP, filiale de Renault qui figure dans la liste des principales entreprises du secteur du « commerce de détail et réparation automobile », n'est pas dans Renault DIF ou que FAURECIA systèmes d'échappement et FAURÉCIA Sièges automobiles ne sont pas dans PSA DAF.

Un consensus a semblé se réaliser sur le fait qu'on puisse diffuser dans les listes d'entreprises les entreprises profilées résultant de profilages de la cible 1.

En réponse à une demande d'explicitation sur le contour de ces entreprises profilées, on ne donnerait pas la composition de chaque entreprise profilée sous forme d'une liste d'unités légales la composant. Par contre, on pourrait communiquer l'ensemble des entreprises profilées d'un groupe, ce qui définit en compréhension le contour des entreprises profilées comme indiqué ci-dessus sur les deux exemples cités, mais pas en extension à partir de la liste des unités légales constituantes⁵. La dénomination de ces entreprises redéfinies devra faire partie des sujets abordés avec le groupe⁶. Pour les cas où des grands groupes auraient refusé le profilage contractualisé et que leur profilage aurait été réalisé par les statisticiens en utilisant l'information diffusée par le groupe, la même règle semble devoir être appliquée puisqu'on aurait utilisé de l'information publique pour réaliser ce profilage.

Pour les profilages réalisés en semi automatique des 2000 groupes de taille intermédiaire, le GT s'est interrogé sur le risque de dévoiler le contour du groupe en diffusant les entreprises profilées. En droit le contour d'un groupe défini par l'Insee à partir de l'enquête LIFI et de l'utilisation complémentaire de DIANE est confidentiel. Toutefois on peut noter que cette information est largement publique. Une proposition a été faite de n'utiliser pour définir les entreprises profilées dans ces profilages semi automatiques que l'information diffusée par le groupe lui-même. En effet la jurisprudence applique le principe qu'il n'y a pas violation du secret statistique si l'information diffusée a par ailleurs déjà été rendue publique par une diffusion par l'entreprise elle-même dans le cadre des obligations de publicité légale ou toute autre diffusion⁷. Ceci permettrait de traiter ces profilages semi automatiques à l'identique des profilages contractualisés pour la diffusion des listes d'entreprises. On devrait ainsi pouvoir continuer à diffuser les 10 principales entreprises du secteur ou de la branche en y incluant les entreprises profilées contractuellement ou en semi automatique.

L'Insee ne juge pas nécessaire de communiquer de manière personnalisée sur leur profilage vers les têtes des petits groupes de la cible 2 profilés automatiquement. Il paraît suffisant d'informer les utilisateurs par une mention dans la méthodologie figurant dans les publications que les résultats présentés portent sur des entreprises reconstituées⁸ et non sur les unités légales de même qu'on indique à l'utilisateur que les non-réponses ont fait l'objet d'une estimation statistique. Toutefois il s'agit d'une indication sur les projets de l'Insee et non d'une position en matière d'application des règles du secret statistique. La jurisprudence actuelle prévoit la possibilité d'une diffusion restreinte aux intéressés de données (contour d'un groupe, en faveur de l'entreprise "tête de groupe" ; appartenance de l'entreprise à un groupe, en faveur de l'entreprise concernée et sauf opposition de la "tête de groupe") collectées par l'INSEE dans l'enquête "Liaisons financières (LIFI)". Cette jurisprudence n'interdirait donc pas de communiquer à l'entreprise tête de groupe quel profilage a été

⁵ On peut ajouter que si la définition des entreprises en compréhension est stable dans le temps il n'en va pas du tout de même de sa composition en unités légales qui bouge assez fréquemment. On a toujours une branche hôtellerie ACCOR mais la liste des hôtels unités légales qui la constitue est fréquemment modifiée.

⁶ C'est bien ce qui s'est produit avec Renault DIF et PSA DAF qui sont des dénominations proposées par les groupes. La plupart du temps ces dénominations se retrouvent sur les sites Internet des groupes.

⁷ On a rediscuté au cours de la réunion du 21/01 pour savoir si une information publiée en dehors d'une obligation légale pouvait être considérée comme « rendue publique ». Toutefois, on voit mal un groupe en appeler à la confidentialité d'une information qu'il aurait lui-même diffusée.

⁸ Il s'agit d'unités d'analyse par opposition aux unités d'observation dont on utilise l'information soit obtenue par enquête ou provenant d'une source administrative pour calculer celle relative à ces unités d'analyse.

retenu pour son groupe⁹ et sauf opposition de la tête de groupe de communiquer à une filiale dans quelle entreprise profilée (c'est à dire en général quel groupe) et dans quel agrégat elle est intégrée. Pour les cas des secteurs et des branches où les 10 principales entreprises contiendraient des entreprises profilées automatiquement, pour pouvoir les faire figurer dans ces listes il conviendrait de vérifier que l'information sur le groupe est publique et non pas connue seulement par la réponse à l'enquête LIFI. Il a été noté que l'algorithme de profilage automatique devrait comporter un module de nommage qui permette de distinguer l'entreprise profilée des unités légales qu'elle regroupe.

Le GT n'a pas explicitement répondu à la question sur la **diffusion des informations autres que d'identification** à savoir :

- les activités secondaires des entreprises
 - les [quatre] variables suivantes :
 - effectifs global d'une entreprise et des établissements (une fois par an) ;
 - catégorie d'importance du chiffre d'affaires ;
 - catégorie d'importance de la part du chiffre d'affaires réalisée à l'exportation ;
 - indicateur de l'exercice d'une activité de recherche.
- qui sont actuellement diffusables dans ces listes.

Il paraît donc souhaitable que le GT exprime un avis sur ce point au cours de la prochaine réunion.

Par ailleurs, la discussion a seulement ébauché la réflexion sur le fait d'adopter éventuellement des règles moins restrictives, notamment sur le contour des entreprises profilées, à des catégories particulières d'utilisateurs : utilisateurs des organisations professionnelles, utilisateurs de l'administration chargés de la politique économique, chargés d'études de l'administration placés sous leur autorité... Elle devra donc être approfondie à la prochaine réunion.

L'habilitation des chercheurs a été ensuite examinée. L'accord est total pour poursuivre ces habilitations qui permettent de communiquer selon la procédure actuelle des informations individuelles provenant des enquêtes et des sources administratives à des chercheurs présentant devant le Comité du Secret Statistique une demande d'accès à des données nécessaires à un projet d'étude. Ces informations comporteraient donc les données individuelles des entreprises profilées. Si le besoin d'en disposer est validé par le Comité du Secret Statistique, on n'exclut pas de fournir également le contour de ces entreprises profilées sous forme d'une liste des filiales regroupées dans une entreprise profilée (extrait du répertoire statistique) si l'étude projetée par le chercheur le nécessite . Il a paru également possible d'habiliter simultanément à des données recueillies auprès d'unités légales et à d'autres données recueillies auprès des entreprises dont certaines issues des profilages lorsqu'un projet de recherche nécessite de mêler ces deux types de données. *Nota du rapporteur : Le statut des données issues de profilages réalisés par les statisticiens eux-mêmes en absence d'accord d'un groupe de la cible 1 a été discuté au cours de la réunion du 21 janvier. Ce débat qui n'avait pas eu lieu le 7 janvier conduit donc à préciser ce § sur l'habilitation des chercheurs dans le compte rendu du 21 janvier.*

La prochaine réunion devra traiter les points complémentaires cités ci-dessus, les habilitations permanentes des différents services statistiques ou d'autres organismes dont la mission statistique nécessite de disposer de l'information individuelle des entreprises et la communication à Eurostat des données agrégées confidentielles et dans certains cas des données individuelles d'entreprise.

Sur les points qui ont déjà été examinés par le GT, le rapporteur proposera pour la prochaine réunion une première rédaction des recommandations que ce GT sera amené à faire au Comité du Secret Statistique du Cnis.

⁹ L'objectif est d'utiliser les entreprises profilées comme unités d'observation pour l'ensemble des enquêtes publiques. Dans le cas où un groupe de la cible 2 est interrogé dans le cadre d'une enquête « entreprise » (par exemple l'enquête CIS, ou TIC,) on devra préciser au répondant le champ sur lequel doit porter sa réponse. On sera donc même obligé dans ce cas de préciser les contours de l'entreprise qui a été redéfinie ou reconstituée à partir des informations contenues dans LIFI ou Diane.

Nombre de groupes par secteur et classes de taille

NAF		US						Total
Division	Libellé	[0;10[[10;20[[20;50[[50 ; 250[[250;5 000[[5000; ∞[
10	Extraction de houille, de lignite et de tourbe	21	3	3	1	1	0	29
11	Extraction d'hydrocarbures ; services annexes	53	4	4	2	0	0	63
12	Extraction de minerais d'uranium	1	0	0	0	0	0	1
13	Extraction de minerais métalliques	43	6	1	1	0	0	51
14	Autres industries extractives	1 116	178	120	41	9	0	1 464
15	Industries alimentaires	56 279	3 553	1 487	723	217	8	62 267
16	Industrie du tabac	2	1	0	1	0	1	5
17	Industrie textile	3 202	442	368	235	56	0	4 303
18	Industrie de l'habillement et des fourrures	6 799	430	328	163	35	0	7 755
19	Industrie du cuir et de la chaussure	1 535	131	104	88	12	1	1 871
20	Travail du bois et fabrication d'articles en bois	7 976	869	550	219	28	0	9 642
21	Industrie du papier et du carton	594	192	167	134	51	1	1 139
22	Edition, imprimerie, reproduction	23 334	1 234	702	334	76	2	25 682
23	Cokéfaction, raffinage, industries nucléaires	22	2	4	4	5	0	37
24	Industrie chimique	1 723	269	284	280	163	9	2 728
25	Industrie du caoutchouc et des plastiques	2 306	622	567	451	139	3	4 088
26	Fabrication d'autres produits minéraux non métalliques	6 653	488	340	186	64	3	7 734
27	Métallurgie	413	117	100	117	50	3	800
28	Travail des métaux	19 283	3 636	2 417	1 044	168	0	26 548
29	Fabrication de machines et d'équipements	11 179	1 322	977	505	156	6	14 145
30	Fabrication de machines de bureau et de matériel informatique	328	30	30	13	3	0	404
31	Fabrication de machines et appareils électriques	2 575	342	259	168	57	6	3 407
32	Fabrication d'équipements de radio, télévision et communication	1 370	251	198	126	40	4	1 989
33	Fabrication d'instruments médicaux, de précision, d'optique et d'horlogerie	9 680	722	401	216	51	1	11 071
34	Industrie automobile	1 127	236	163	120	59	6	1 711
35	Fabrication d'autres matériels de transport	2 424	136	99	78	31	5	2 773
36	Fabrication de meubles ; industries diverses	24 576	778	501	263	50	0	26 168
37	Récupération	3 200	270	159	65	8	0	3 702
40	Production et distribution d'électricité, de gaz et de chaleur	1 845	14	29	11	9	3	1 911
41	Captage, traitement et distribution d'eau	123	10	8	9	4	0	154
45	Construction	330 422	16 361	7 412	1 667	150	8	356 020
50	Commerce et réparation automobile	67 936	3 743	1 382	606	124	3	73 794
51	Commerce de gros et intermédiaires du commerce	122 747	7 770	4 568	1 805	343	11	137 244
52	Commerce de détail et réparation d'articles domestiques	374 378	8 639	4 081	1 688	313	18	389 117
55	Hôtels et restaurants	206 911	7 770	2 902	669	95	11	218 358
60	Transports terrestres	68 701	3 657	2 472	917	167	5	75 919
61	Transports par eau	1 483	53	21	16	6	0	1 579
62	Transports aériens	252	13	16	8	9	1	299

63	Services auxiliaires des transports	7 143	781	535	296	109	6	8 870
64	Postes et télécommunications	3 946	161	101	30	15	4	4 257
70	Activités immobilières	340 430	1 978	772	450	87	5	343 722
71	Location sans opérateur	21 203	363	183	100	27	1	21 877
72	Activités informatiques	35 271	1 596	1 051	534	129	9	38 590
73	Recherche et développement	2 170	128	141	70	21	0	2 530
74	Services fournis principalement aux entreprises	340 029	11 635	6 181	2 764	506	26	361 141
80	Education	33 452	960	478	129	20	0	35 039
85	Santé et action sociale	317 893	3 725	2 292	848	156	8	324 922
90	Assainissement, voirie et gestion des déchets	1 177	126	87	40	14	1	1 445
92	Activités récréatives, culturelles et sportives	163 249	1 618	993	385	59	4	166 308
93	Services personnels	111 818	1 477	347	94	13	1	113 750
95	Activités des ménages en tant qu'employeur de personnel domestique	21	0	0	0	0	0	21
Total		2 740 414	88 842	46 385	18 714	3 905	184	2 898 444
	Nombre de secrets dans chaque colonne	2	2	1	4	1	9	1
	Nombre de cases non nulles diffusables dans chaque colonne	49	47	47	45	45	24	50
	Nombre de cases nulles dans chaque colonne	0	2	3	2	5	18	0
	Nombre de cases dans chaque colonne	51	51	51	51	51	51	51