

# Le code statistique non signifiant (CSNS)

un nouveau service pour faciliter  
les appariements

Séminaire CNIS du 28 janvier 2022



# 01 QU'EST CE QUE LE CSNS ?

## L'offre de service

- Proposer une « aide aux appariements de données individuelles »
  - accroître l'usage des sources administratives et d'enquêtes
  - enrichir les statistiques sans augmenter la charge d'enquêtes
  - pour le service statistique public (SSP)
- Produire une clé d'appariement
  - non signifiante
  - sécurisée
  - pérenne sur 10 ans, avec processus de renouvellement
- Définir une organisation pour utiliser le CSNS au sein du SSP
  - processus organisationnel d'accès au service, avec conventionnement
  - applicatif sécurisé et dédié pour l'ensemble des utilisateurs du SSP

## Les fondements juridiques du projet CSNS

- Origine du projet : la loi pour une République numérique de 2016 – article 34
  - introduction de la notion de Code statistique non signifiant
  - décret n°2016-1930 qui donne le cadre d'utilisation du CSNS
  - arrêté du 28 septembre 2020 sur l'opération cryptographique, la sécurité des clés de chiffrement et les conditions de stockage et renouvellement
- Article 30 de la loi Informatique et libertés
- Règlement général sur la protection des données, notamment le principe de minimisation des données

## Produit final

- A chaque individu (identifié) est associé un CSNS
- Calculé :
  - **directement à partir du NIR (hachage puis cryptage)**
    - *service ouvert depuis octobre 2021*
  - **ou à partir de traits d'identité (nom, prénoms, date et lieu de naissance) permettant de déterminer le NIR**
    - implique une étape préalable d'identification statistique au RNIPP**
      - *service ouvert au 2ème trimestre 2022*
- **Un exemple de CSNS**  
v1:3zUYVUpJFX9g7z3oi3zKSIKXB0yLiE4oGmxsQi1Z2atWI0n8IfnmVrxWUiJqeKHHTvcH+i0PvuO2991M

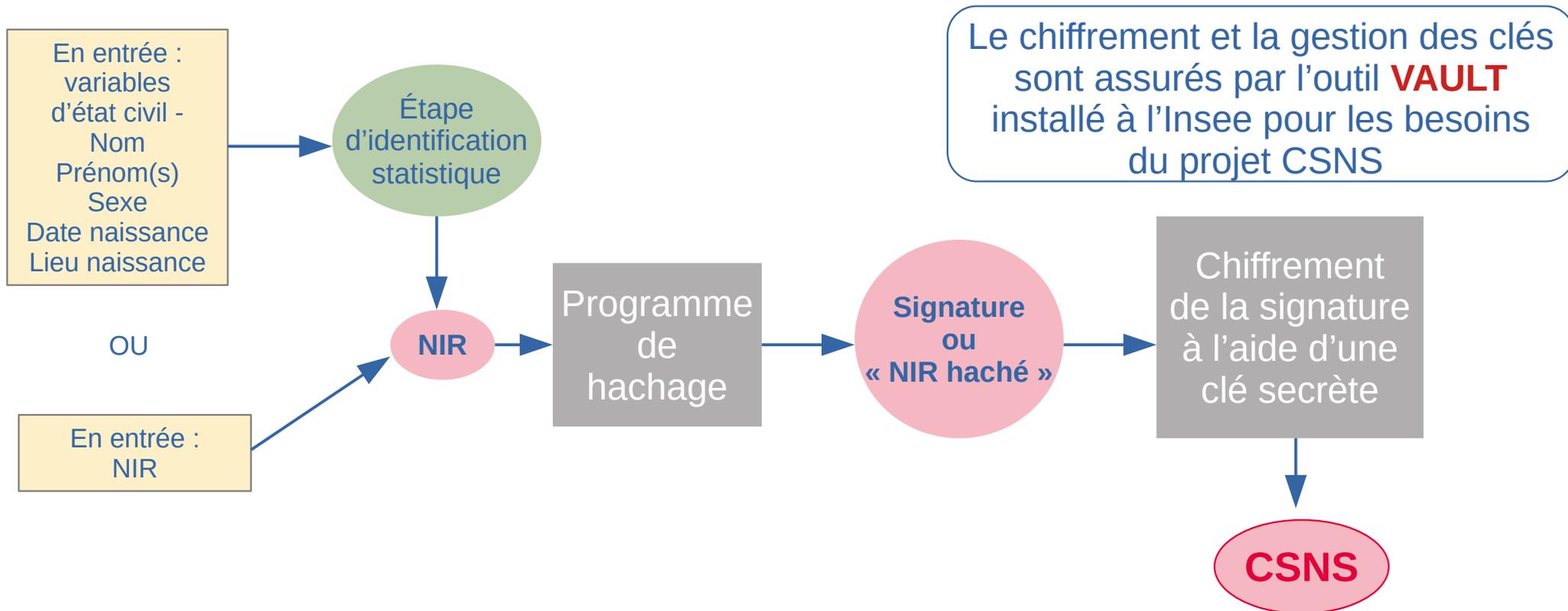
# 02

## COMMENT ÇA MARCHE ?

## Les étapes d'appariement

1. Signer la convention de sous-traitance avec l'Insee pour un service statistique ministériel (ou la Charte d'usage en interne pour un service de l'Insee)
2. En cas d'appariement entre fichiers de producteurs différents, établir les règles de collaboration (convention)
3. Déclarer le traitement
4. Inscrire le traitement dans le programme de travail transmis au CNIS
5. Demander le CSNS à l'Insee (application dédiée)
6. Réaliser l'appariement selon les modalités prévues avec les partenaires
7. Conserver les données selon des règles précises (durée et isolement du CSNS)

## Comment est pratiquement obtenu le CSNS ?



## Exemple d'appariement sur la base du CSNS

Propriétaire du fichier A

**Fichier A**

```
01 var1 var2 ... varN
02 var1 var2 ... varN
03 var1 var2 ... varN
....
zz var1 var2 ... varN
```

**Table de CSNS associée au fichier A**

```
01 CSNSa indQual*
02 CSNSa indQual
03 CSNSa indQual
....
zz CSNSa indQual
```

Chaque propriétaire demande un CSNS pour son fichier

Propriétaire du fichier B  
Responsable de traitement

**Fichier B**

```
11 var1' var2' ... varN'
12 var1' var2' ... varN'
13 var1' var2' ... varN'
....
1z var1' var2' ... varN'
```

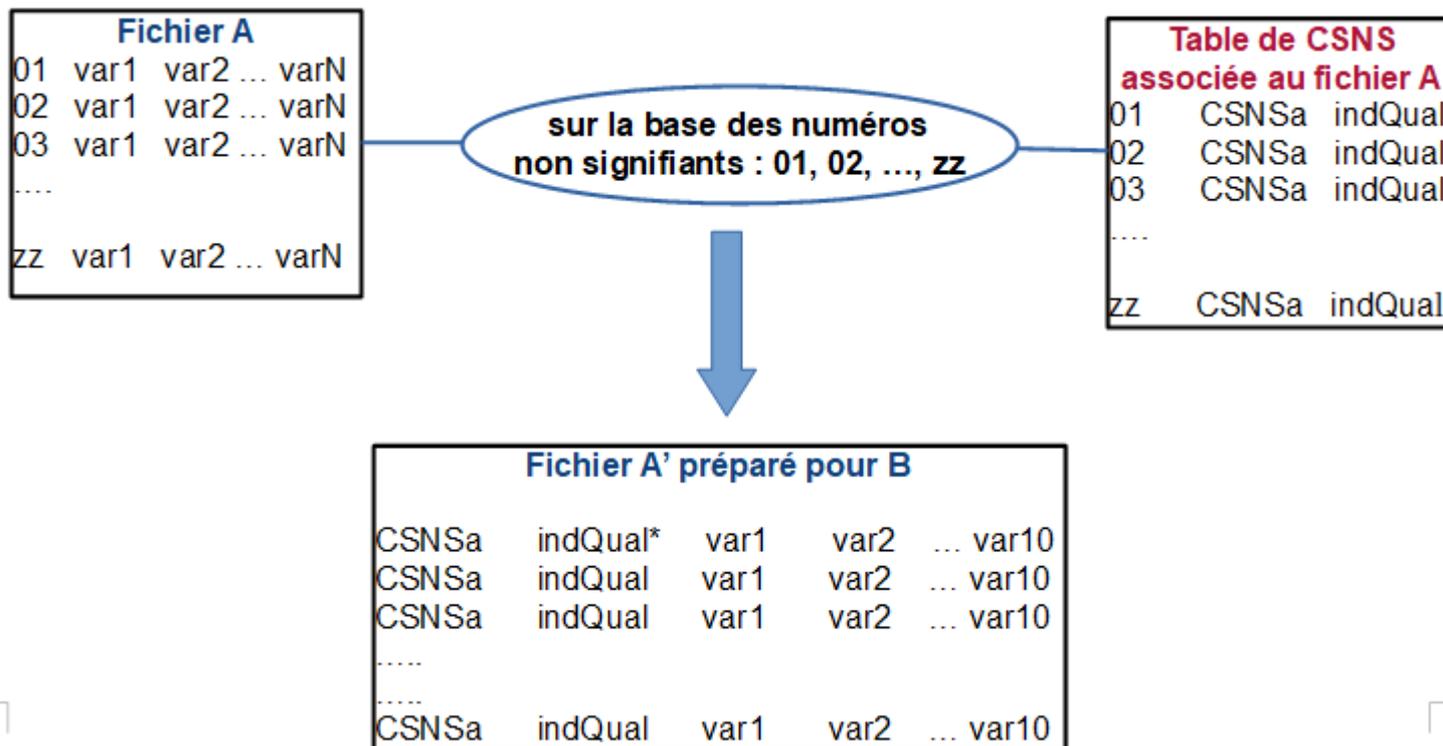
**Table de CSNS associée au fichier B**

```
11 CSNSb indQual*
12 CSNSb indQual
13 CSNSb indQual
....
1z CSNSb indQual
```

Tables de passage contenant les CSNS à conserver de façon hermétique aux autres données

## Exemple

### Étape 1 - Préparation du fichier A' pour le responsable B



## Exemple

### Étape 2 - Utilisation du fichier A' pour le rapprochement avec B

#### Rapprochement de A et B sur la base du CSNS

var1	var2	...	var10	indQual*a	CSNSa	=	CSNSb	indQual*b	01'
var1	var2	...	var10	indQual*a	CSNSa	=	CSNSb	indQual*b	02'
var1	var2	...	var10	indQual*a	CSNSa	=	CSNSb	indQual*b	03'
.....									
.....									
.....									
var1	var2	...	var10	indQual*a	CSNSa	=	CSNSb	indQual*b	tt'
var1	var2	...	var10	indQual*a	CSNSa				
.....									
var1	var2	...	var10	indQual*a	CSNSa		CSNSb	indQual*b	xx'
							.....		
							CSNSb	indQual*b	zz'

Bleu : apparié

Jaune : dans A mais pas dans B (au sens du CSNS)

Rose : dans B mais pas dans A

## Exemple

### Étape 3 – Résultat de l'enrichissement du fichier B par A

#### Appariement sur la base du numéro d'ordre (fichier B)

var1	var2	...	var10	indQual*a	<b>01'</b>	indQual*b	var1'	var2'	...	varN'
var1	var2	...	var10	indQual*a	<b>02'</b>	indQual*b	var1'	var2'	...	varN'
var1	var2	...	var10	indQual*a	<b>03'</b>	indQual*b	var1'	var2'	...	varN'
....	....	....	....	....	....	....	....	....	....	....
var1	var2	...	var10	indQual*a	<b>tt'</b>	indQual*b	var1'	var2'	...	varN'
						<b>xx'</b>	indQual*b	var1'	var2'	... varN'
						.....				
						<b>zz'</b>	indQual*b	var1'	var2'	... varN'

Bleu : apparié

Rose : dans B mais pas dans A

## Conservation du CSNS – sécurisation

- **Le service de l’Insee délivrant le CSNS ne conserve aucune donnée** (*dès que le traitement est validé par le demandeur*)
  - destruction du fichier en entrée (NIR ou données état civil)
  - destruction du fichier de résultats (CSNS)
  
- **Le CSNS est conservé par le demandeur**
  - à l’écart des autres variables (statistiques ou identifiantes)
  - dans une table contenant uniquement le CSNS et un numéro « repère » pour chaque ligne, hermétique aux autres données
  - ne pas conserver le CSNS au-delà du strict nécessaire

## En cours d'élaboration

- **Objectif de trouver le juste équilibre entre identifier le maximum de personnes et limiter le risque d'erreur**
- **Fournir des indicateurs qualité**
- **Test en cours avec 4 services statistiques ministériels (Drees, Dares, SDES, SIES)**

## Les grands principes retenus – 3 étapes

- recherche d'appariement exact avec le RNIPP
- recherche simple avec relâche de variable une à une et conservation du résultat uniquement si un seul candidat
- recherche plus élaborée avec relâche sur plusieurs variables simultanément

## Contacts

[insee.fr](https://www.insee.fr)



Lionel Espinasse – [lionel.espinasse@insee.fr](mailto:lionel.espinasse@insee.fr)

Séverine Gilles – [severine.gilles@insee.fr](mailto:severine.gilles@insee.fr)

Département de la démographie (DSDS)