



Les répertoires

Pivots indispensables et méconnus des SI

Pascal Rivière, Inspection générale de l'Insee

FNRS
« Analyse critique et
amélioration de la
qualité de l'information
numérique »
17 mai 2023

Plan

Des mots à clarifier

Le contenu d'un répertoire

Un système vivant

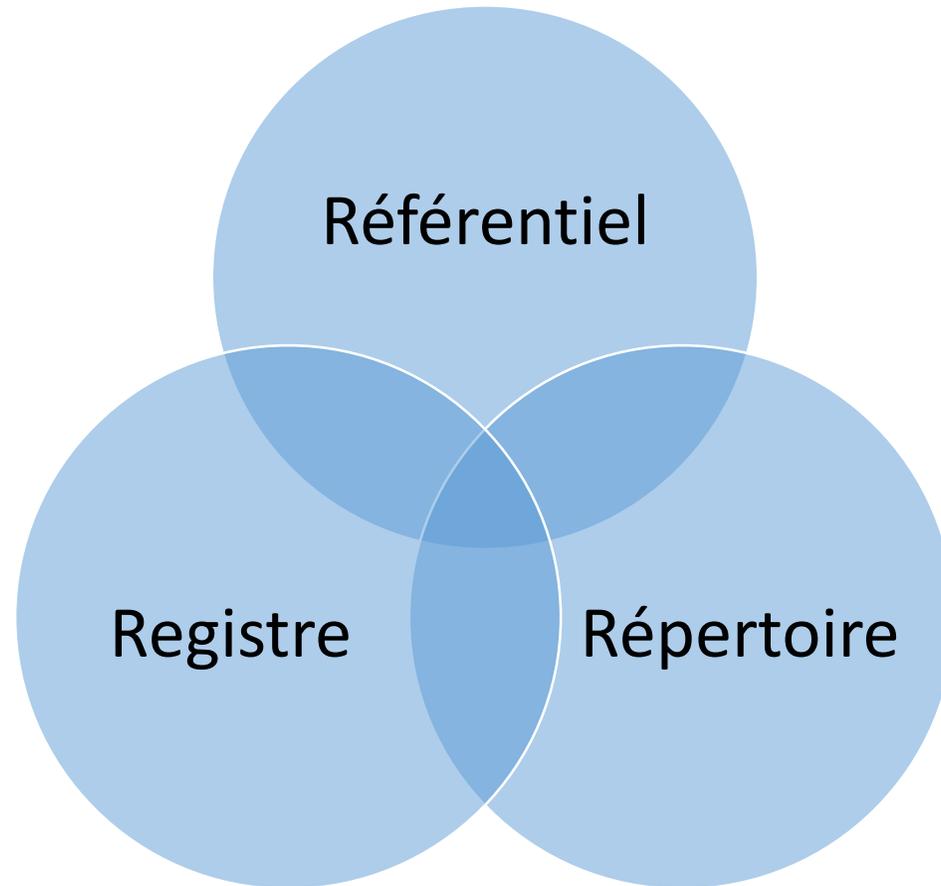
Cinq propriétés fondamentales

Les différentes dimensions de la qualité

Environnement et évolution

Un enchevêtrement de termes proches

- Mais aussi :
 - Inventaire
 - Recensement
 - Et les termes anglais : directory, repository, register, ...



Répertoire

- *« Recueil, livre comportant des données classées selon un certain ordre : Consulter le répertoire des fournisseurs d'une entreprise. Ensemble des oeuvres qu'a l'habitude d'interpréter un artiste, une troupe. Ensemble des grands ballets classiques et romantiques. Ensemble des instructions de commande d'un ordinateur. »*
- Répertorier, cataloguer
- Usage spécifique dans divers domaines : artistique, linguistique
- Idée de collection d'objets, d'ensemble défini en extension
- Plus précisément, **ensemble d'occurrences d'une même entité.**

Registre

- *« Cahier sur lequel on note des faits, des noms, des chiffres dont on veut garder le souvenir... (Petit Robert 2006). Livre où l'on inscrit les faits, les choses dont on veut garder le souvenir. (Larrousse). Dès le XIIIe siècle : livres où l'on écrit les actes, les affaires de chaque jour. Dès le XVIe siècle on parle de « tenir registre de quelque chose », c'est-à-dire de consigner régulièrement quelque chose, d'en effectuer le recueil. »*
- Registres paroissiaux dès le XIVe (naissances, sépultures), registres de transactions commerciales dès l'antiquité : événements datés
- Stockage des informations, mais surtout le garant de leur intégrité
- De manière générale, un registre a deux fonctions principales :
 - **Conserver des informations de manière permanente**
 - Assurer l'**inaltérabilité** des données qui y sont présentes.

Référentiel

- « *Ensemble structuré d'informations ou encore un système de référence lié à un champ de connaissance, notamment en vue d'une pratique ou d'une étude, et dans lequel se trouvent des éléments de définitions, de solutions, de pratiques ou autres sujets relatifs de ce champ de connaissance* »
- Référentiel signifie « qui a trait à la référence ».
 - Latin *referre*, littéralement rapporter, voire attribuer. Référé, référence, référencer, référencement, référendaire, référendum, référent
- Nombreux domaines : physique, linguistique, psychologie, formation
- Repère au sens large, source d'information officielle, langage commun et reconnu, cadre sur lequel on peut fonder son action
 - Exemple de l'histoire du système métrique
- Trois idées : ce à quoi on se **réfère**, **langue commune**, **facilitateur d'échanges**

Contexte d'un système d'information

- Référentiels :
 - Répertoires : agents, fournisseurs, clients
 - Nomenclatures, classifications : produits, activités
- Apparaissent clairement dans la cartographie fonctionnelle d'un SI
- La dimension « enregistrement » va de soi (base de données)
- Facilitateur d'échanges car utilisation par plusieurs directions métier

Ce dont on va parler ...

Répertoires à *statut référentiel* au sein d'un
système d'information

Exemples de répertoires, en vrac

- Individus
- Formations
- Professionnels de santé
- Véhicules
- Etablissements d'enseignement
- Etudiants
- Carrières
- Entreprises
- Electeurs
- Etablissements de santé
- Equipements
- Fournisseurs
- Agents
- Clients

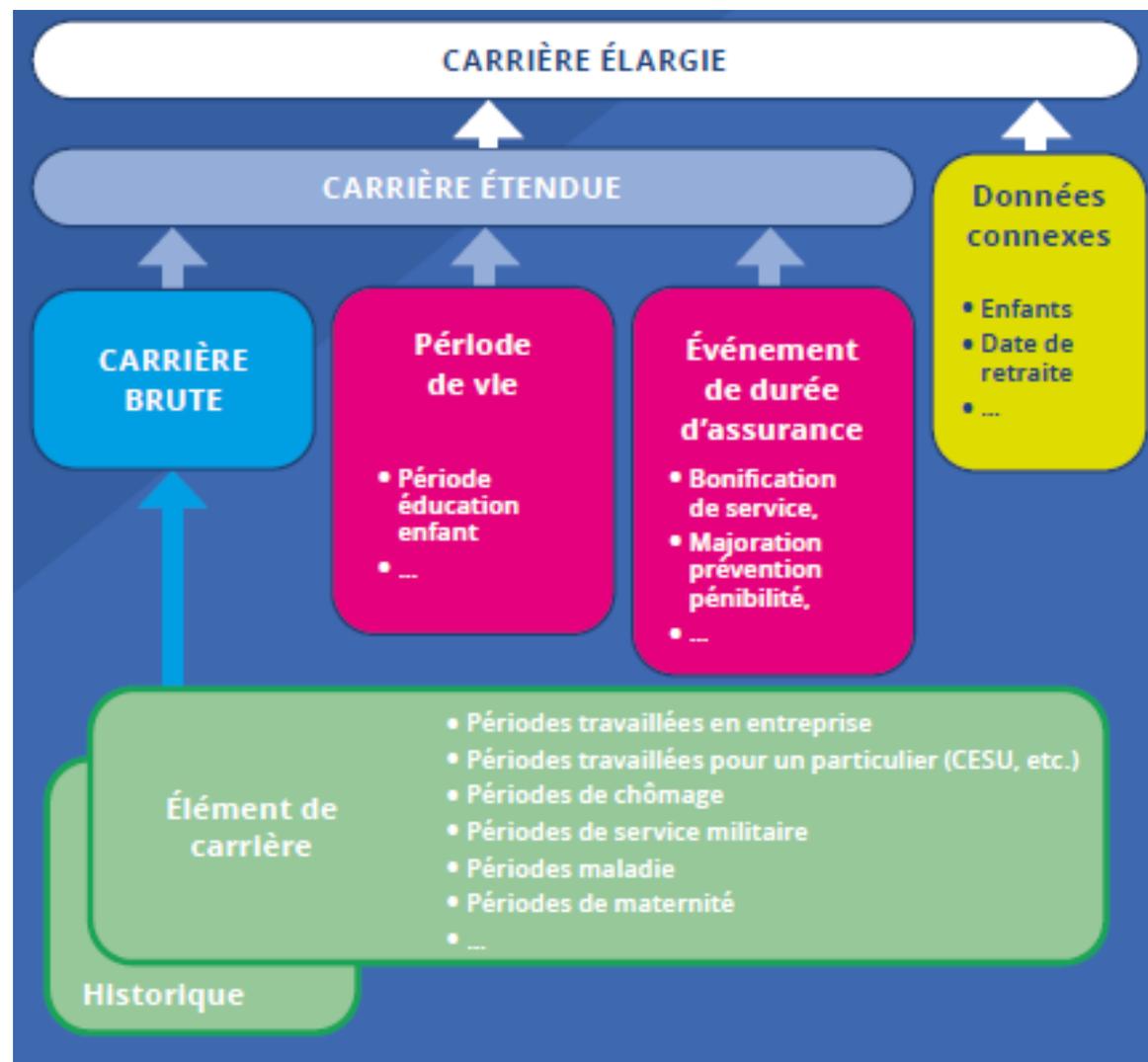
Le contenu



Trois dimensions à distinguer

- Ensemble de référence : l'ensemble en intension que l'on est censé représenter de façon exhaustive
 - Ex. entreprises actives ayant une activité marchande et présentes sur le territoire
- L'objet, ou entité
 - Pas toujours simple : client/prospect, hôpital
- Les variables
 - Toutes les caractéristiques de l'objet qu'on va conserver
- Banal dans un système d'information ? Oui et non

Exemple d'objet complexe : la carrière



Les variables possibles

- Identifiant
- Traits d'identité
- Dénominations
- Localisation
- Catégorisation, simple ou complexe
- Dates (de création, de cessation, ...)
- Variables de gestion
- Liens vers d'autres objets
- Historique

Exemple : répertoire d'établissements d'enseignement

0950926R - Collège Stendhal - Ouvert le 21/10/1969 Données mises à jour le 19/12/2022																												
i Identification	<table><tr><td>Commune</td><td>95250 - FOSSES</td></tr><tr><td>Secteur</td><td>PU - Public</td></tr><tr><td>Ministère principal</td><td>06 - MINISTERE DE L'EDUCATION NATIONALE</td></tr><tr><td>Ministère secondaire</td><td></td></tr><tr><td>Nature</td><td>340 - COLLEGE</td></tr><tr><td>Type</td><td>CLG - COLLEGE</td></tr><tr><td>Contrat</td><td>99 - SANS OBJET</td></tr></table>	Commune	95250 - FOSSES	Secteur	PU - Public	Ministère principal	06 - MINISTERE DE L'EDUCATION NATIONALE	Ministère secondaire		Nature	340 - COLLEGE	Type	CLG - COLLEGE	Contrat	99 - SANS OBJET	🏠 État <table><tr><td>Etat</td><td>OUVERT</td></tr><tr><td>Date d'ouverture administrative</td><td>21/10/1969</td></tr><tr><td>Date de rentrée</td><td>21/10/1969</td></tr><tr><td>Motif d'ouverture</td><td>Ex-Nihilo</td></tr><tr><td>Motif de fermeture</td><td></td></tr><tr><td>Date de fermeture</td><td></td></tr></table>	Etat	OUVERT	Date d'ouverture administrative	21/10/1969	Date de rentrée	21/10/1969	Motif d'ouverture	Ex-Nihilo	Motif de fermeture		Date de fermeture	
Commune	95250 - FOSSES																											
Secteur	PU - Public																											
Ministère principal	06 - MINISTERE DE L'EDUCATION NATIONALE																											
Ministère secondaire																												
Nature	340 - COLLEGE																											
Type	CLG - COLLEGE																											
Contrat	99 - SANS OBJET																											
Etat	OUVERT																											
Date d'ouverture administrative	21/10/1969																											
Date de rentrée	21/10/1969																											
Motif d'ouverture	Ex-Nihilo																											
Motif de fermeture																												
Date de fermeture																												
📄 Appellations	<table><tr><td>Sigle</td><td>CLG</td></tr><tr><td>Dénomination principale</td><td>COLLEGE</td></tr><tr><td>Dénomination complémentaire (patronyme)</td><td>STENDHAL</td></tr><tr><td>Appellation officielle</td><td>Collège Stendhal</td></tr><tr><td>Appellation d'usage</td><td></td></tr><tr><td>Appellation d'édition</td><td>STENDHAL CHEMIN DE BEAUMONT</td></tr></table>	Sigle	CLG	Dénomination principale	COLLEGE	Dénomination complémentaire (patronyme)	STENDHAL	Appellation officielle	Collège Stendhal	Appellation d'usage		Appellation d'édition	STENDHAL CHEMIN DE BEAUMONT	✉ Adresse <table><tr><td>Type de voie</td><td>chemin</td></tr><tr><td>Libellé de la voie</td><td>de Beaumont</td></tr><tr><td>Traduction postale</td><td>Chemin de Beaumont</td></tr><tr><td>Code postal</td><td>95470</td></tr><tr><td>Localité d'acheminement</td><td>FOSSES</td></tr></table>	Type de voie	chemin	Libellé de la voie	de Beaumont	Traduction postale	Chemin de Beaumont	Code postal	95470	Localité d'acheminement	FOSSES				
Sigle	CLG																											
Dénomination principale	COLLEGE																											
Dénomination complémentaire (patronyme)	STENDHAL																											
Appellation officielle	Collège Stendhal																											
Appellation d'usage																												
Appellation d'édition	STENDHAL CHEMIN DE BEAUMONT																											
Type de voie	chemin																											
Libellé de la voie	de Beaumont																											
Traduction postale	Chemin de Beaumont																											
Code postal	95470																											
Localité d'acheminement	FOSSES																											
🏛 Administration	<table><tr><td>Hébergement</td><td>12 - SANS INTERNAT AVEC DEMI PENSION</td></tr><tr><td>Catégorie juridique</td><td>200 - ETABL PUBLIC LOCAL D ENSEIGNEMENT</td></tr><tr><td>Catégorie financière</td><td>4 - CATEGORIE FINANCIERE 4</td></tr><tr><td>Situation comptable</td><td>3 - RATTACHE A UNE AGENCE COMPTABLE</td></tr><tr><td>SIRET</td><td>19950926600011</td></tr><tr><td>FINES</td><td></td></tr><tr><td>APE</td><td>8531Z</td></tr></table>	Hébergement	12 - SANS INTERNAT AVEC DEMI PENSION	Catégorie juridique	200 - ETABL PUBLIC LOCAL D ENSEIGNEMENT	Catégorie financière	4 - CATEGORIE FINANCIERE 4	Situation comptable	3 - RATTACHE A UNE AGENCE COMPTABLE	SIRET	19950926600011	FINES		APE	8531Z	👤 Groupements <table><tr><td>Tête de groupement</td><td>27 - POLE INCLUSIF D'ACCOMPAGNEMENT LOCALISE - 22 adhérente(s)</td></tr><tr><td>Adhérente au groupement</td><td>27 - POLE INCLUSIF D'ACCOMPAGNEMENT LOCALISE - 0950926R 42 - GROUPEMENT COMPTABLE - 0950647M 43 - GRETA - 0952078T 44 - CENTRE D INFORMATION ET D ORIENTATION - 0950976V</td></tr></table>	Tête de groupement	27 - POLE INCLUSIF D'ACCOMPAGNEMENT LOCALISE - 22 adhérente(s)	Adhérente au groupement	27 - POLE INCLUSIF D'ACCOMPAGNEMENT LOCALISE - 0950926R 42 - GROUPEMENT COMPTABLE - 0950647M 43 - GRETA - 0952078T 44 - CENTRE D INFORMATION ET D ORIENTATION - 0950976V								
Hébergement	12 - SANS INTERNAT AVEC DEMI PENSION																											
Catégorie juridique	200 - ETABL PUBLIC LOCAL D ENSEIGNEMENT																											
Catégorie financière	4 - CATEGORIE FINANCIERE 4																											
Situation comptable	3 - RATTACHE A UNE AGENCE COMPTABLE																											
SIRET	19950926600011																											
FINES																												
APE	8531Z																											
Tête de groupement	27 - POLE INCLUSIF D'ACCOMPAGNEMENT LOCALISE - 22 adhérente(s)																											
Adhérente au groupement	27 - POLE INCLUSIF D'ACCOMPAGNEMENT LOCALISE - 0950926R 42 - GROUPEMENT COMPTABLE - 0950647M 43 - GRETA - 0952078T 44 - CENTRE D INFORMATION ET D ORIENTATION - 0950976V																											
🎓 Spécificités	<table><tr><td>SPE - PRESENCE CLASSE ENSEIGNT SPECIAL (CLG)</td></tr></table>	SPE - PRESENCE CLASSE ENSEIGNT SPECIAL (CLG)																										
SPE - PRESENCE CLASSE ENSEIGNT SPECIAL (CLG)																												
🔗 Rattachements	<table><tr><td>UAI fille(s)</td><td>390 - SEGPA - SEGPA - (FI) - 0951152L</td></tr></table>	UAI fille(s)	390 - SEGPA - SEGPA - (FI) - 0951152L																									
UAI fille(s)	390 - SEGPA - SEGPA - (FI) - 0951152L																											
📍 Zones éducation	<table><tr><td>22 - SECTEUR SCOLAIRE - 0950807 - FOSSES</td></tr><tr><td>📍 23 - DISTRICT SCOLAIRE - 09508 - SARCELLES</td></tr><tr><td>24 - BASSIN DE FORMATION - 25021 - 095D-GONESSE</td></tr></table>	22 - SECTEUR SCOLAIRE - 0950807 - FOSSES	📍 23 - DISTRICT SCOLAIRE - 09508 - SARCELLES	24 - BASSIN DE FORMATION - 25021 - 095D-GONESSE																								
22 - SECTEUR SCOLAIRE - 0950807 - FOSSES																												
📍 23 - DISTRICT SCOLAIRE - 09508 - SARCELLES																												
24 - BASSIN DE FORMATION - 25021 - 095D-GONESSE																												
	📍 Zones géographiques <table><tr><td>Agglomération urbaine</td><td>95401 - FOSSES</td></tr><tr><td>Zone d'emploi</td><td>0056 - ROISSY - SUD PICARDIE</td></tr><tr><td>Canton</td><td>A9509 - FOSSES</td></tr><tr><td>Arrondissement</td><td>952 - SARCELLES</td></tr><tr><td>Département</td><td>095 - VAL-D'OISE</td></tr><tr><td>Académie</td><td>25 - VERSAILLES</td></tr><tr><td>Base de gestion</td><td>25 - VERSAILLES</td></tr><tr><td>Région</td><td>11 - ILE-DE-FRANCE</td></tr><tr><td>Pays</td><td>100 - FRANCE</td></tr></table>	Agglomération urbaine	95401 - FOSSES	Zone d'emploi	0056 - ROISSY - SUD PICARDIE	Canton	A9509 - FOSSES	Arrondissement	952 - SARCELLES	Département	095 - VAL-D'OISE	Académie	25 - VERSAILLES	Base de gestion	25 - VERSAILLES	Région	11 - ILE-DE-FRANCE	Pays	100 - FRANCE									
Agglomération urbaine	95401 - FOSSES																											
Zone d'emploi	0056 - ROISSY - SUD PICARDIE																											
Canton	A9509 - FOSSES																											
Arrondissement	952 - SARCELLES																											
Département	095 - VAL-D'OISE																											
Académie	25 - VERSAILLES																											
Base de gestion	25 - VERSAILLES																											
Région	11 - ILE-DE-FRANCE																											
Pays	100 - FRANCE																											
	📍 Géolocalisation <table><tr><td>Coordonnée X</td><td>664239.7</td></tr><tr><td>Coordonnée Y</td><td>6889398.0</td></tr><tr><td>Qualité appariement</td><td>11 - Correcte</td></tr><tr><td>Qualité localisation</td><td>21 - Rue</td></tr><tr><td>Date de géolocalisation</td><td>22/11/2022</td></tr><tr><td>Source</td><td>90 - IGN</td></tr></table>	Coordonnée X	664239.7	Coordonnée Y	6889398.0	Qualité appariement	11 - Correcte	Qualité localisation	21 - Rue	Date de géolocalisation	22/11/2022	Source	90 - IGN															
Coordonnée X	664239.7																											
Coordonnée Y	6889398.0																											
Qualité appariement	11 - Correcte																											
Qualité localisation	21 - Rue																											
Date de géolocalisation	22/11/2022																											
Source	90 - IGN																											

Répertoire socle ou adossé

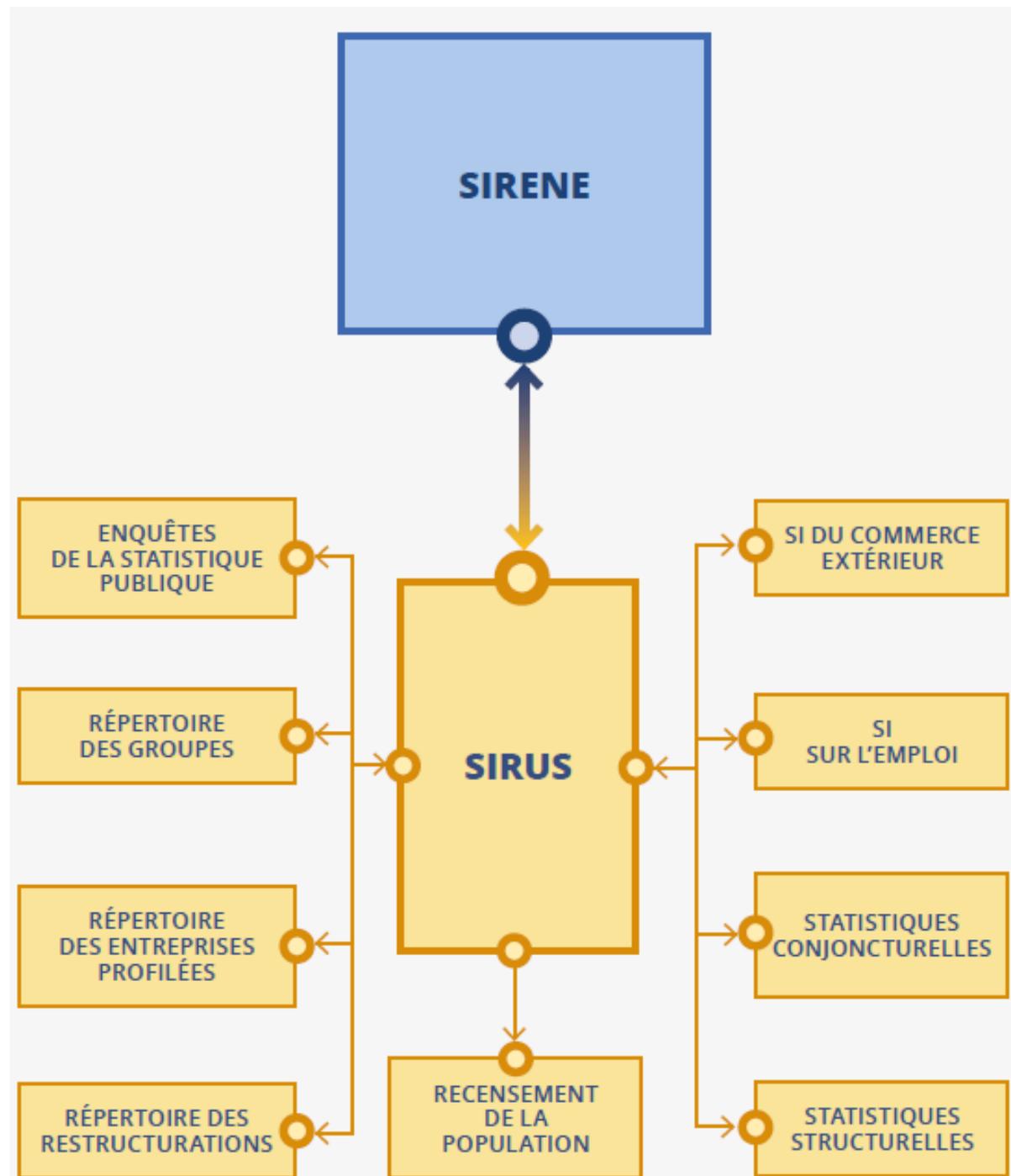
- Socle : assure le lien entre traits d'identité et identifiant
 - Donc processus d'immatriculation, qui crée l'objet
 - Algorithme d'identification
 - Ex. répertoires d'individus, d'entreprises, d'agents
- Adossé : s'appuie sur un répertoire-socle
 - Souvent associé à l'idée de « rôle »
 - Ex. répertoires de clients, d'électeurs, ...
 - Variables associées à ce rôle, par ex. données de contact pour un client

Répertoire administratif / statistique

- Administratif : conséquences individuelles en gestion
 - « Retour de bâton » immédiat : ex. conséquence sur prestations sociales
 - Donc besoin de mises à jour régulières en fonction d'événements
 - ... peu de marge de manœuvre
 - Ex. répertoires d'entreprises, d'électeurs, d'individus
- Statistique : usage à des fins statistiques, aucun impact d'une décision sur un individu
 - Besoin de mise à jour : fonction de l'usage
 - Acceptation de moindre « preuve » pour meilleur qualité statistique
 - Ex. à l'Insee répertoire Sirius d'unités statistiques entreprises, mais aussi Base permanente des équipements
 - Ex. des « signes de vie »

Exemple

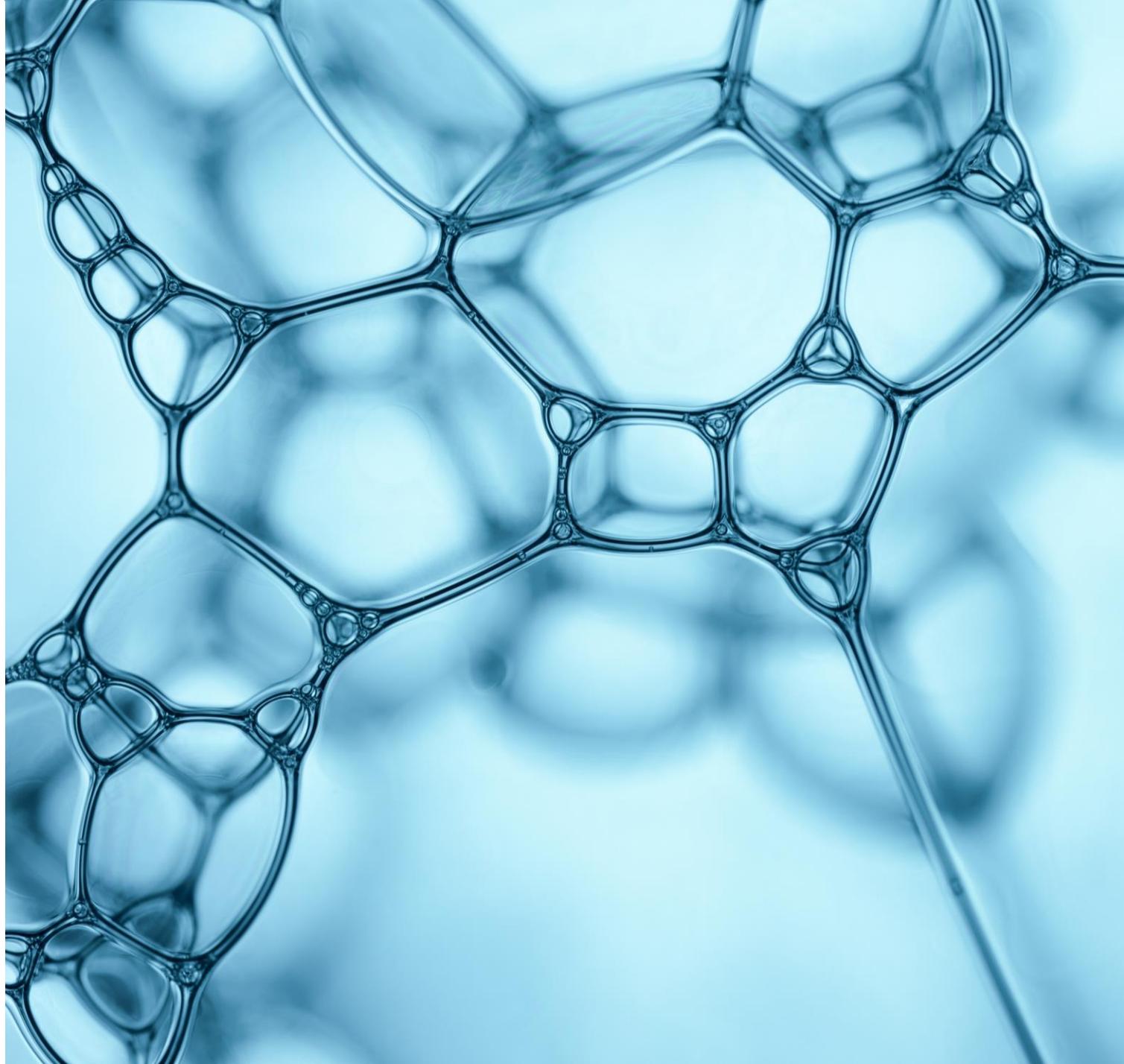
- Sirius : répertoire statistique adossé au répertoire administratif Sirene
- Interne à la statistique publique
- Différence entre « unités économiques » dans Sirius et « unités légales » dans Sirene
- Meilleur contrôle du champ



Remarques supplémentaires

- Données de gestion : dans le répertoire mais à statut non référentiel
 - Données de contact, nom du gestionnaire, numéro de dossier, etc.
- Autres variables non référentielles : idéalement à éviter
 - Ex. données budgétaires d'établissements d'enseignement
- Vision 360° : en réalité très éloigné de la logique répertoire
 - Construction ad hoc

Un système
vivant

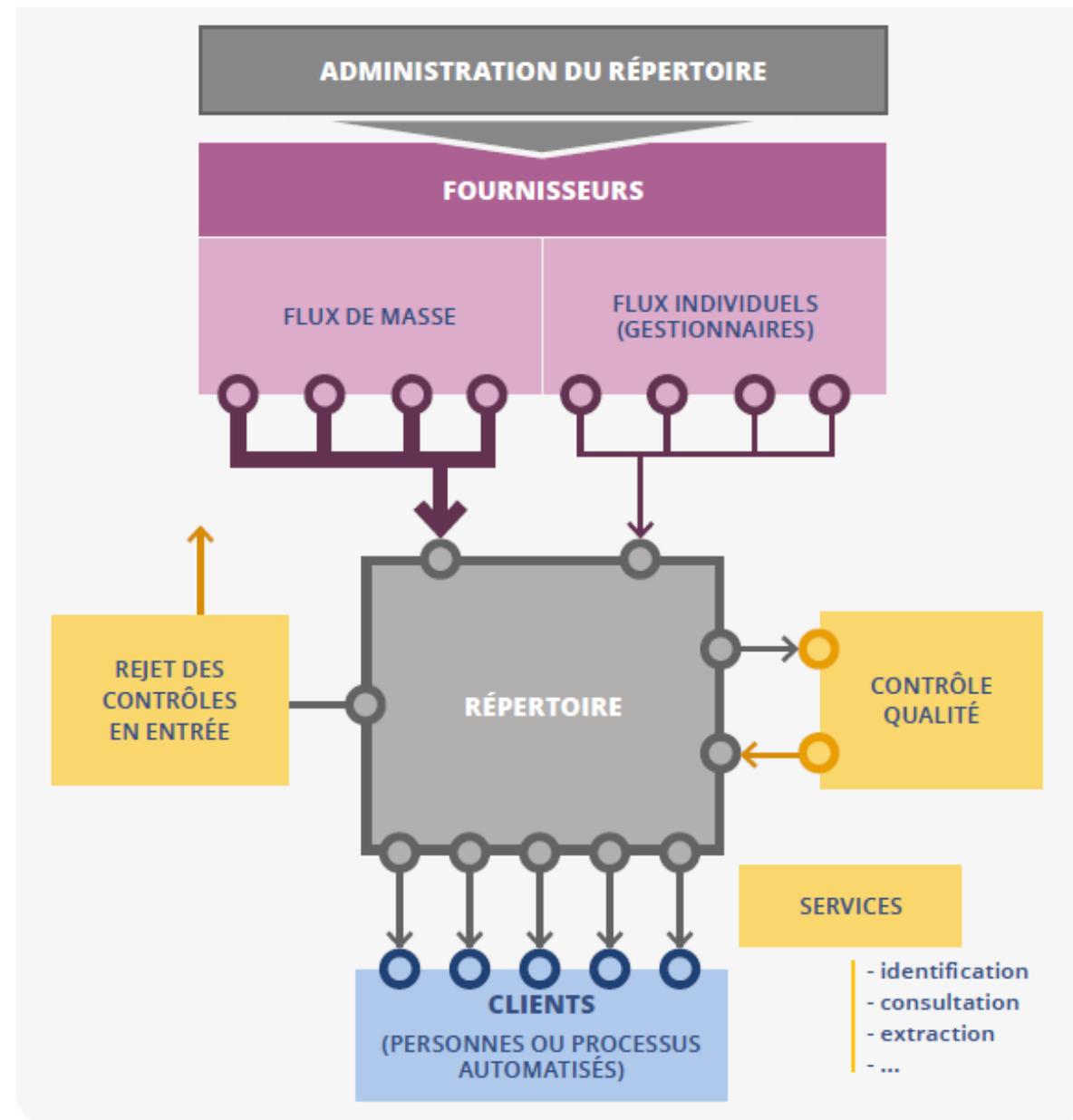


Représentation imagée d'un répertoire

- Image traditionnelle apportée par la dimension référentielle : ce seraient des données « **gravées dans le marbre** »
→ NON !
- Image plus appropriée d'un cœur, d'une « **pompe aspirante-refoulante d'informations** », et d'un véritable système associé
- Ex. pour un répertoire d'agents : données de recrutement, de promotion, ... en entrée, et en aval, usage pour la paie, les congés, les affectations, ...

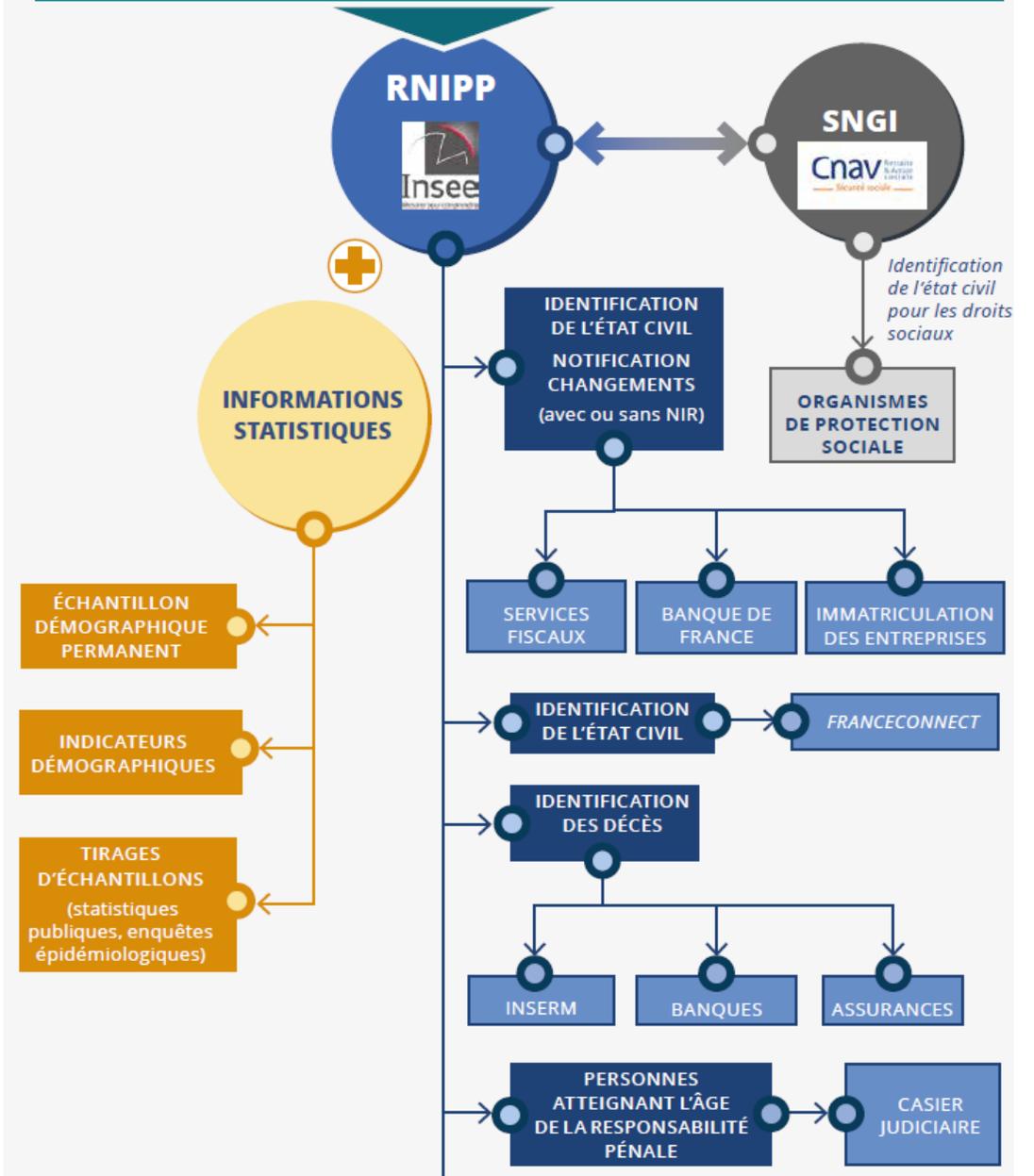
Le « système » répertoire

- Des flux d'entrée : l'alimentation du répertoire
 - Individuelle ou de masse
- La base de données elle-même
- Des flux sortants : les services rendus par le répertoire
- La régulation de l'ensemble : contrôle qualité et administration du répertoire



Exemple :
le Répertoire national
d'identification
des personnes physiques
(RNIPP)

35 000 COMMUNES
EN FRANCE
+ COLLECTIVITÉS D'OUTRE-MER
HORS NOUVELLE-CALÉDONIE



La problématique de l'alimentation du répertoire

- Quels fournisseurs ?
 - Ex. organismes habilités à mettre à jour une carrière
- Quelles sources légitimes ?
 - Nécessité de support juridique, ex. répertoire d'établissements de santé Finess
 - La question des pièces justificatives est un sujet
- Alimentation de masse / individuelle
 - Mise à jour par des gestionnaires → organisation du travail, formation
- Contrôles de validité
 - Notion de norme d'échange
- Caractérisation de la cinématique des flux
- Problème de la méconnaissance des événements
 - Dissymétrie création - cessation

Sources du RNIPP

Les bulletins d'état civil sont accessibles sur le site de l'Insee.fr.
(<https://www.insee.fr/fr/information/1303477>)

NAISSANCE

Formule de déclaration de naissance (Bulletin de Naissance) en français. Le formulaire est divisé en plusieurs sections pour recueillir les informations relatives à l'enfant, à la mère, au père, et à l'acte de naissance. Il comprend des champs pour les dates, les lieux, les noms, et des cases à cocher pour indiquer des détails spécifiques.

MARIAGE

Formule de déclaration de mariage (Bulletin de Mariage) en français. Le formulaire recueille les informations sur les deux époux, le lieu et la date du mariage, et les conditions de l'union. Il inclut des sections pour les noms, les professions, et les détails de l'acte.

CONCLUSION D'UN PACS

Formule de déclaration de conclusion d'un PACS (Bulletin de Solidarité PACS) en français. Ce formulaire est utilisé pour enregistrer la conclusion d'un Pacte Civil de Solidarité. Il recueille les noms, les professions, et les dates des partenaires.

DÉCÈS

Formule de déclaration de décès (Bulletin de Décès) en français. Le formulaire recueille les informations sur le défunt, le lieu et la date du décès, et les détails de l'acte. Il comprend des sections pour les noms, les professions, et les circonstances du décès.

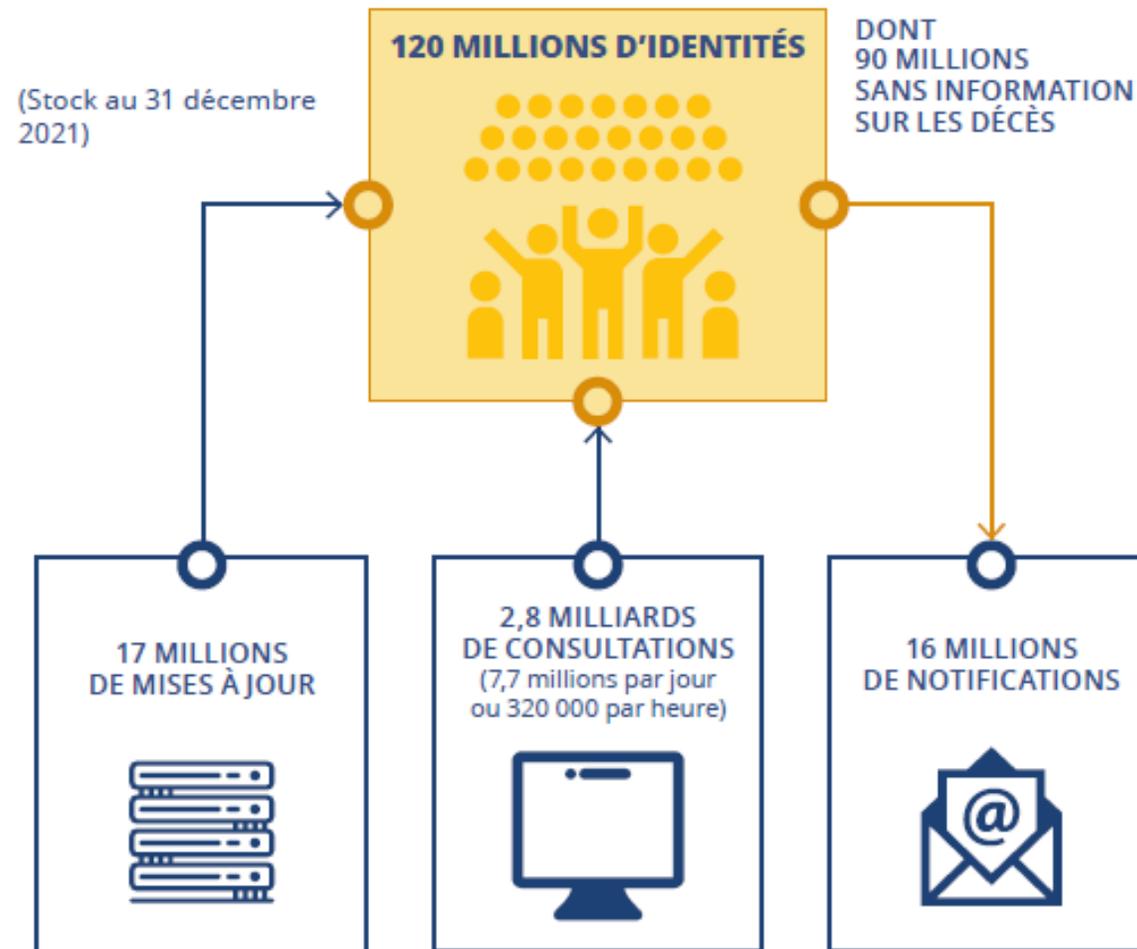
MAIS AUSSI :

- Transcription relative à un jugement déclaratif de naissance
- Bulletin d'enfant sans vie
- Transcription relative à un jugement d'adoption plénière
- Transcription relative à un jugement déclaratif de décès ou d'absence

Services rendus par un répertoire

- Processus d'immatriculation (répertoire-socle)
- Identification (répertoire-socle)
 - Indispensable pour « retrouver » l'entité, ou vérifier son existence
 - Essentiel pour les activités d'appariement
 - Attention à la distinction traits d'identité - empreinte
- Consultation
- Extraction
 - Risque d'aspiration du répertoire par recopies
 - Double temporalité : ex. liste d'entreprises au 31/12... vue au 17/5
- Exposition via API

Illustration avec le SNGI (Système National de Gestion des Identifiants)



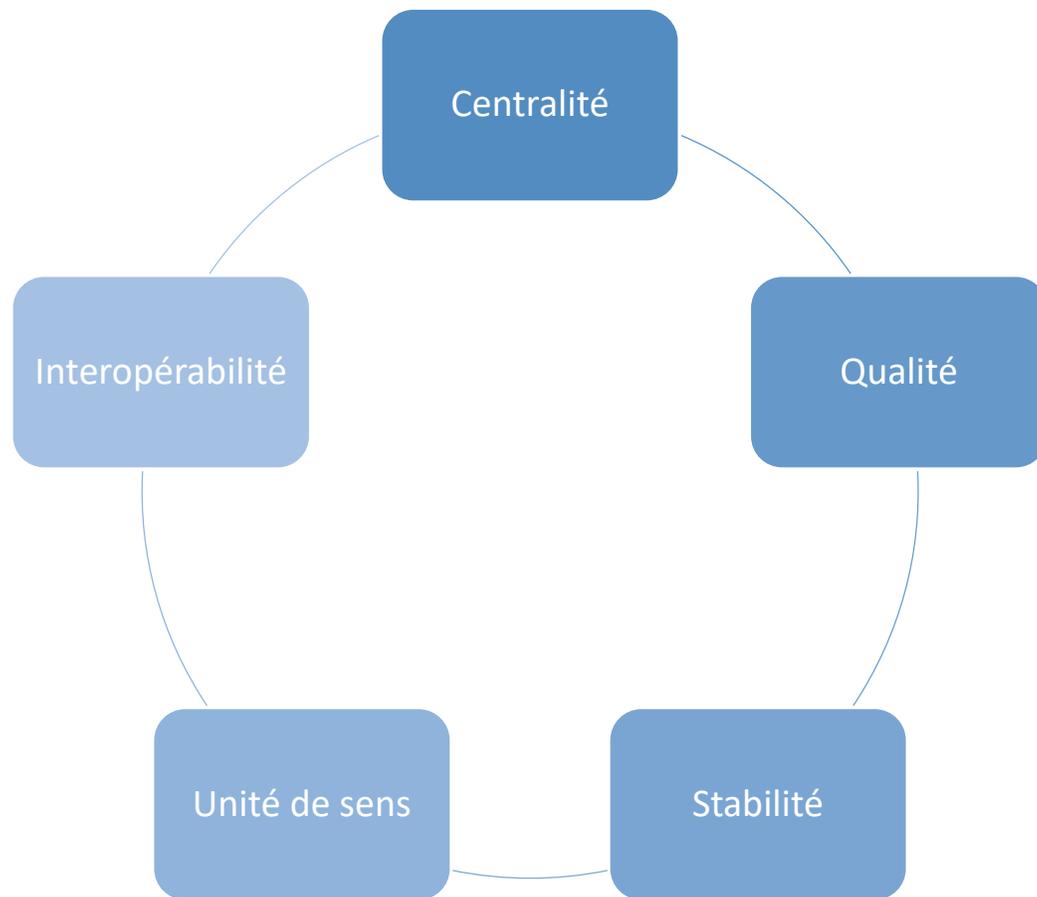
Temporalités : deux niveaux différents

- Temporalité courte de modification des données
 - Ex. plus de 10000 créations ou modifications par jour dans Sirene
- Temporalité longue de la modification du système
 - Nouvelles variables : une seule variable nouvelle dans un répertoire, c'est très exigeant
 - Ex. ajout de l'appartenance à un groupe
 - Nouveaux types d'objets
 - Ex. nouveaux types d'établissements de santé
 - Mises à jour des contrôles
 - Avec démarche de type ATMS
 - Nouveaux fournisseurs de données
 - Donc mécanisme d'habilitation
 - Cas du répertoire de carrières RGCU

Caractéristiques d'un répertoire



Cinq propriétés



Centralité

- Légitimité permettant au répertoire de jouer un rôle central pour un ensemble d'acteurs
- Cas d'une organisation très centralisée : répertoire d'établissements d'enseignement Ramsese
- Cas d'un support juridique fort : Rnipp, Finess, Sirene ...
- Attention : centralité
 - *Relativement à un certain domaine (cf. l'ensemble en intension)*
 - *Vis-à-vis d'un certain nombre d'acteurs (cf. différence RNIPP – SNGI)*

Stabilité

- Intuitivement, des données « de référence » ne vont pas bouger en permanence
- Différence entre fréquence de mise à jour « macro » et fréquence de mise à jour « individuelle »
 - Enormément de mises à jour dans les grands répertoires, mais très peu au niveau individuel (ex. peu d'événements de mises à jour Rnipp)
- Fréquence de mise à jour : pas le bon indicateur de stabilité
- Principale sujet : découplage répertoire – processus opérationnels
 - Exemples : nombre de lits occupés, achats d'un client

Unité de sens

- Renvoie au besoin de « lisibilité » du répertoire et à la cohérence de son évolution future
- L'unité de sens porte d'abord sur les objets
 - Contre-exemples : demande d'immatriculation de ruches et éoliennes
- Elle porte aussi sur les attributs des objets
 - Eviter le caractère fourre-tout sinon les demandes des utilisateurs vont se multiplier
 - Cohérence vis-à-vis du « rôle » (ex. nombre d'enfants pour un électeur)
 - Ex déjà cité des établissements d'enseignement
- De vrais risques, plus particulièrement si le répertoire est un succès

Interopérabilité

- Technique (performance, disponibilité, sécurité)
 - Rien de vraiment spécifique à un répertoire
- Syntaxique
 - Formats utilisés : XML, JSON, ...
 - Rien de spécifique pour un répertoire
- Sémantique
 - Identifiants, concepts, nomenclatures, ...
 - Lié à l'environnement dans lequel gravite le répertoire
 - Là en revanche exigence très forte et spécifique pour un répertoire

Qualité d'un
répertoire



Des spécificités en termes de qualité

- A quoi comparer ?
- Notions de surcouverture, sous-couverture
 - Surcouverture : la question des « faux actifs » (doublons, plus classique) et toujours la dissymétrie création/cessation
 - Sous-couverture : très difficile car pas de repérage automatique
- Exigences de qualité ++
 - Car usages multiples et variés
- Importance de la fraîcheur car impacts en gestion
- Compromis sur les définitions et concepts en raison de la variété des usages
 - Ex. usages statistique et administratif, avec par ex. la notion de « signe de vie »

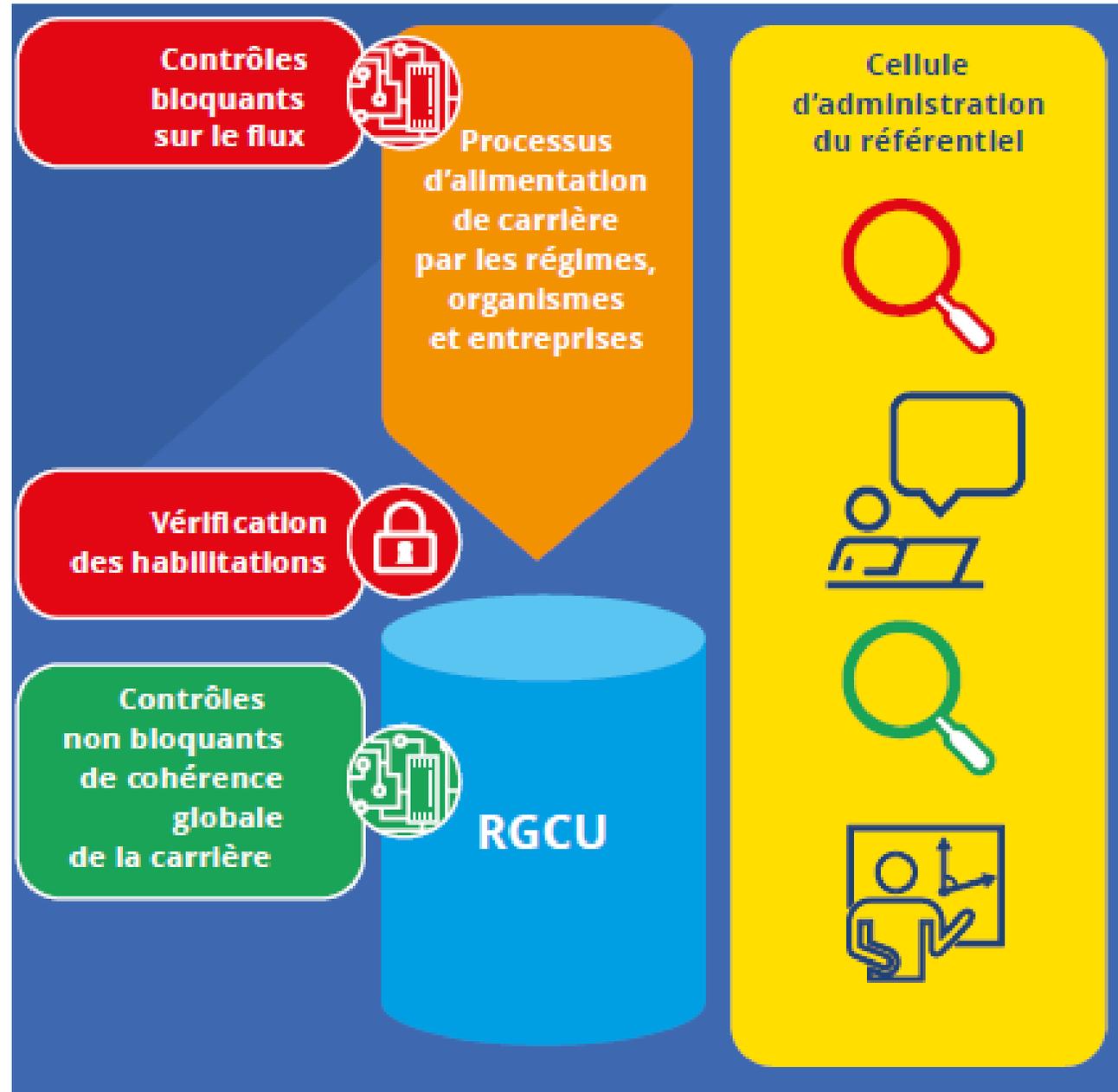
Qualité de l'armature / conception

- Rigueur de la modélisation
 - Caractérisation de l'ensemble de référence
 - Exigence de qualité élevée pour le modèle conceptuel de données, sur la signification des concepts (par ex. qu'est-ce-qu'une carrière ?)
 - Attention particulière aux métadonnées (de date notamment)
- Description détaillée de la cinématique des flux
 - Que se passe-t-il si le flux ne « passe » pas les contrôles, quelles étapes
- Travail sur le support juridique
- Nécessité de communiquer et partager sur cette conception

Qualité des données entrantes

- Préalable : contrôle d'habilitation des fournisseurs de données (y compris gestionnaires de répertoire)
- Normes d'échange pour le contrôle du contenu du flux transmis
 - Normes EDI-CFE (Sirene), A (SNGI), R (RGCU)
- Contrôle de cohérence entre flux et base mise à jour (existence de modifications impossibles)
 - Ex. Incohérence d'éléments de carrière dans le RGCU
- Contrôles manuels supplémentaires
 - Ex. adresses, vérifications d'existence en cas de doute
- Contrôle de pièces justificatives
 - Ex validité du passeport pour les nés à l'étranger dans le SNGI
- Eventuellement, contrôles macro
 - Suivi mensuel du nombre total de mises à jour, une évolution très importante de ce nombre peut cacher un problème

Cas du répertoire de carrières : vue d'ensemble

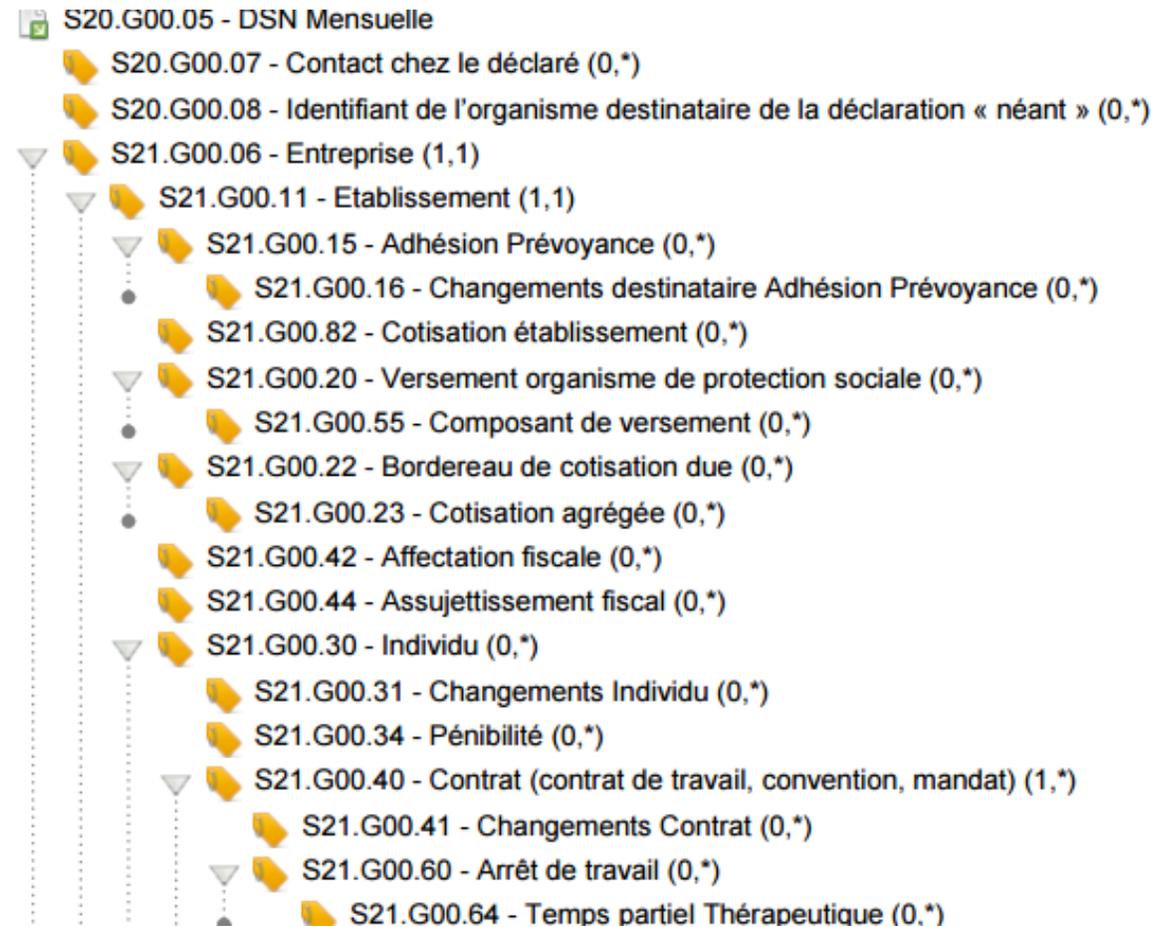
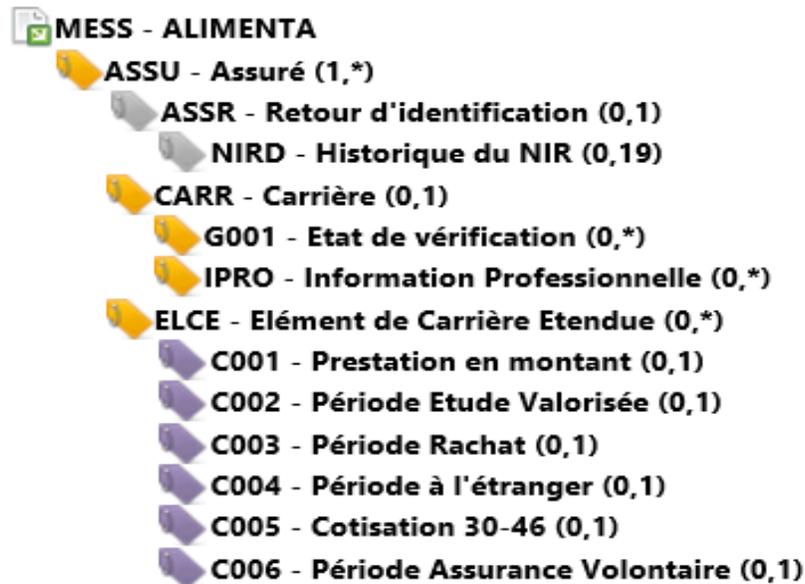


Notion de norme d'échanges

- Modèle de message : règles à respecter par le message transmis
 - Structure du message : contrôles d'existence, de cardinalité, d'ordre
 - Contrôles unaires
 - Typage : numérique, date, chaîne de caractères, ...
 - Vérification d'appartenance à une liste de valeurs
 - Vérification de structure de la donnée (expression régulière)
 - Contrôles de cohérence entre rubriques
 - Si telle rubrique est présente alors telle autre est absente
 - Relations mathématiques entre valeurs (ex. date début \leq date fin)
- Modèle de données explicité (en principe ...)
- Cinématique des flux
- Documentation technique complexe (centaines de pages)
- Versioning rigoureux et en cohérence (par ex. nomenclatures)

La notion de norme d'échange

- Extrait de la structure en blocs fonctionnels d'un flux entrant dans le RGCU



- Extrait de la structure en blocs fonctionnels d'une déclaration sociale nominative

La notion de norme d'échange

- Exemple d'extrait de liste de valeurs possibles pour une rubrique



☰ X 📄 [3,3]



112 - invalidité catégorie 1
114 - invalidité catégorie 2
116 - invalidité catégorie 3
200 - COP (congrés payés)
301 - congé de formation professionnelle
501 - congé divers non rémunéré
507 - chômage intempéries
601 - mobilité volontaire sécurisée
602 - chômage sans rupture de contrat
603 - détention provisoire

- Exemple de règle de contrôle inter-rubriques (RGCU)



CCH-11 : Si un bloc Versement cumulé (A003) est présent, alors la rubrique Type de Élément de carrière étendue (ELCE.001) doit être différent de '01', '02', '03', '04', '05', '06', '07', '08', '10', '11', '12', '13', '14', '15', '16', '17', '18', '19', '20', '21', '22', '23', '24', '25', '26', '27', '28', '29', '30', '31', '32', '33', '34', '35', '36' ou '37'.

Gouvernance

- Instances de gouvernance du répertoire
 - Contenant, ou pas, les instances de gouvernance de la norme
 - Comités stratégiques ou plus techniques (ex. pour validation documentation)
- Organismes membres des instances en majorité déterminés au départ
- Nature des décisions à prendre
 - Extension du champ, des concepts
 - Nouveaux services rendus
 - Nouvelles variables (souvent en lien avec évolutions réglementaires)
 - Evolution des contrôles, domaines, ... d'ordre plus technique
- Nécessairement des compromis et des frustrations
- Travail facilité par des tableaux de bord
 - Suivi quantitatif des consultations, des modifications

Qualité de service

- Dimension informatique : pas très spécifique
 - Performance, disponibilité, sécurité ... : exigences non fonctionnelles
 - S'applique aux services de consultation, identification, extraction, ...
- Service dans un sens plus général : rôle des gestionnaires
 - Rapidité de prise en compte des demandes
 - Réponse aux contestations
 - Accompagnement, rôle explicatif
 - Qualité de la documentation, du site, ...
- Exemples
 - à l'Insee : rôle des gestionnaires Sirene, et des « pôles Insee-Contacts
 - CNAV : rôle des gestionnaires du Sandia pour le SNGI (répertoire d'individus de la protection sociale)

Environnement et évolutions

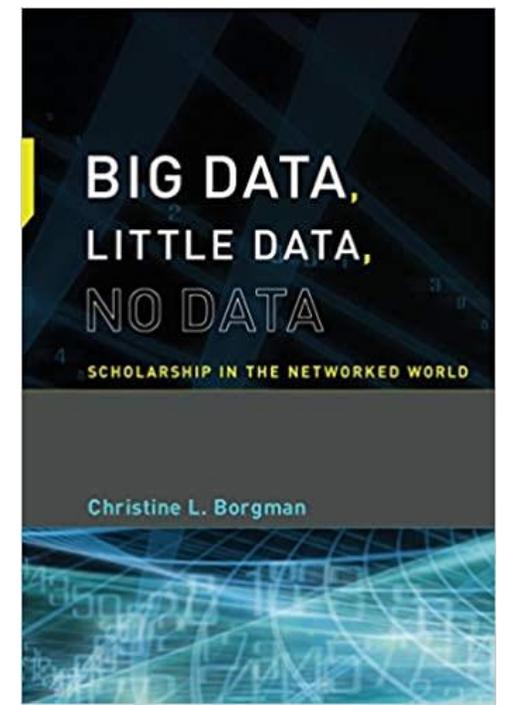
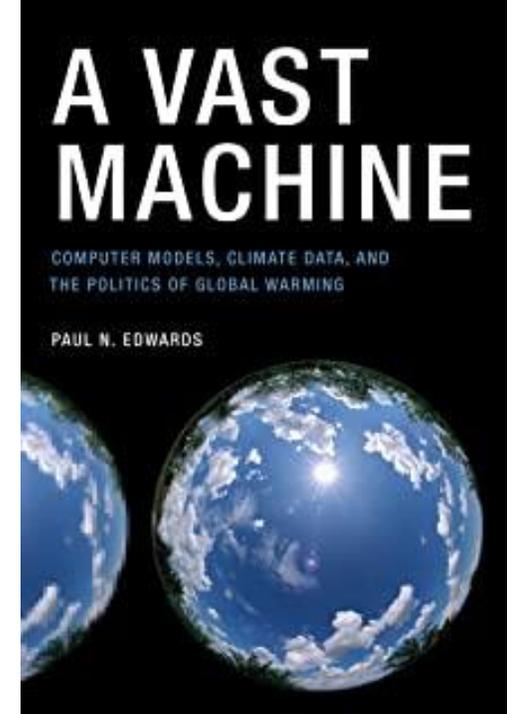


Environnements normés

- Les répertoires font souvent partie d'un environnement référentiel plus large comprenant :
 - Des normes d'échanges
 - Des nomenclatures (ex. nomenclature d'activités pour Sirene)
 - Des définitions
 - Des méthodes
 - ... le tout avec un versioning précis
- Exemples
 - DmFa en Belgique pour la Banque-Carrefour
 - Environnement associé à la Déclaration sociale nominative
 - Agence du numérique en santé, Agence technique de l'information hospitalière
 - Sandre pour le système d'information de l'eau
 - SIMM (milieux marins)

Echanges de données dans le domaine scientifique

- Plus généralement : notion d'infrastructure de connaissances (ce qui inclut aussi les processus d'acquisition de données)
 - Christine Borgman : Big data, Little data, No data (2015)
- Mise en évidence de l'importance de ces infrastructures pour que les données soient comparables dans l'espace et dans le temps
 - Paul Edwards : A vast machine (2011) → climatologie
- Essentiel dans le contexte de l'open data et de l'e-administration
 - Renvoie à l'importance des conventions, au caractère construit des données
- Les répertoires ne sont ici qu'un aspect parmi d'autres des éléments de référence nécessaires



Projet répertoire

- Projets ambitieux, de grande ampleur, et qui posent un problème de portage
 - Ex du répertoire des unités d'une organisation
- Temps important de conception, de consultation des utilisateurs
- Multiples projets dans le projet : reprise de données, transition, bascule, norme d'échange, interface du gestionnaire, administration de répertoire, GED- pièces justificatives, ...
- Tensions car intérêts divergents
- Sous-estimation systématique charge, délai, complexité

Evolution à long terme

- Rigueur originelle : pas une garantie absolue
- Risque de relâchement sur le niveau d'exigence ... sauf à lancer un nouveau projet, une « v2 »
- Nouveaux acteurs pouvant avoir :
 - Une vision simpliste du sujet
 - Une moindre sensibilisation à la cohérence d'ensemble
 - Une attente de réponse rapide à des besoins
- ... et des problèmes peuvent survenir ...

Sujets de recherche possibles ?

- Conception d'algorithmes d'identification
- Modélisation des situations de « faux actifs »
- Optimisation de l'architecture des contrôles
- Cadrage théorique de la notion de norme d'échanges
- Analyse d'impact de la qualité du répertoire sur des usages ciblés
- Travaux de nature plus sociologique à mener sur le fonctionnement des instances de gouvernance
- De façon plus abstraite, il y a certainement une réflexion à mener autour des types d' « objet » et leurs liens avec la qualité des données

Merci de votre attention

pascal.riviere@insee.fr