

# La norme OAIS: un modèle conceptuel et fonctionnel pour la pérennisation des données numériques

**Lourdes FUENTES-HASHIMOTO**  
**ministère des Affaires étrangères, Archives**  
**diplomatiques**

**Marie LAPERDRIX**  
**Archives nationales**

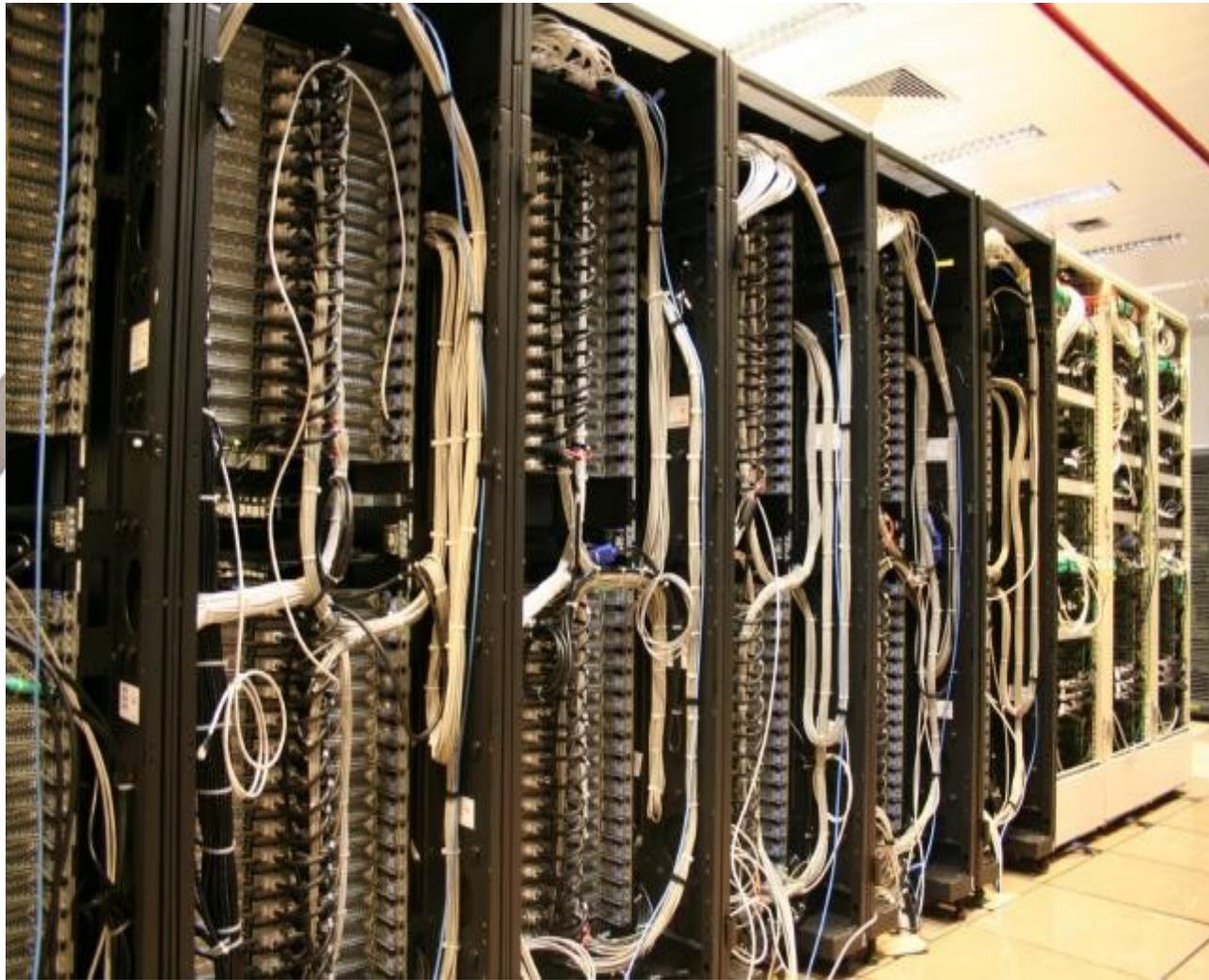
Qu'est-ce que c'est?



Est-ce que j'ai un ordinateur avec un lecteur de disquettes?

Est-ce que j'ai le logiciel nécessaire à lire l'information?

Est-ce que l'information est fiable?  
Vais-je relire ce qui était écrit à l'origine?



**La durée de vie du matériel informatique de stockage est inférieure à la durée de vie souhaitée pour les données numériques**

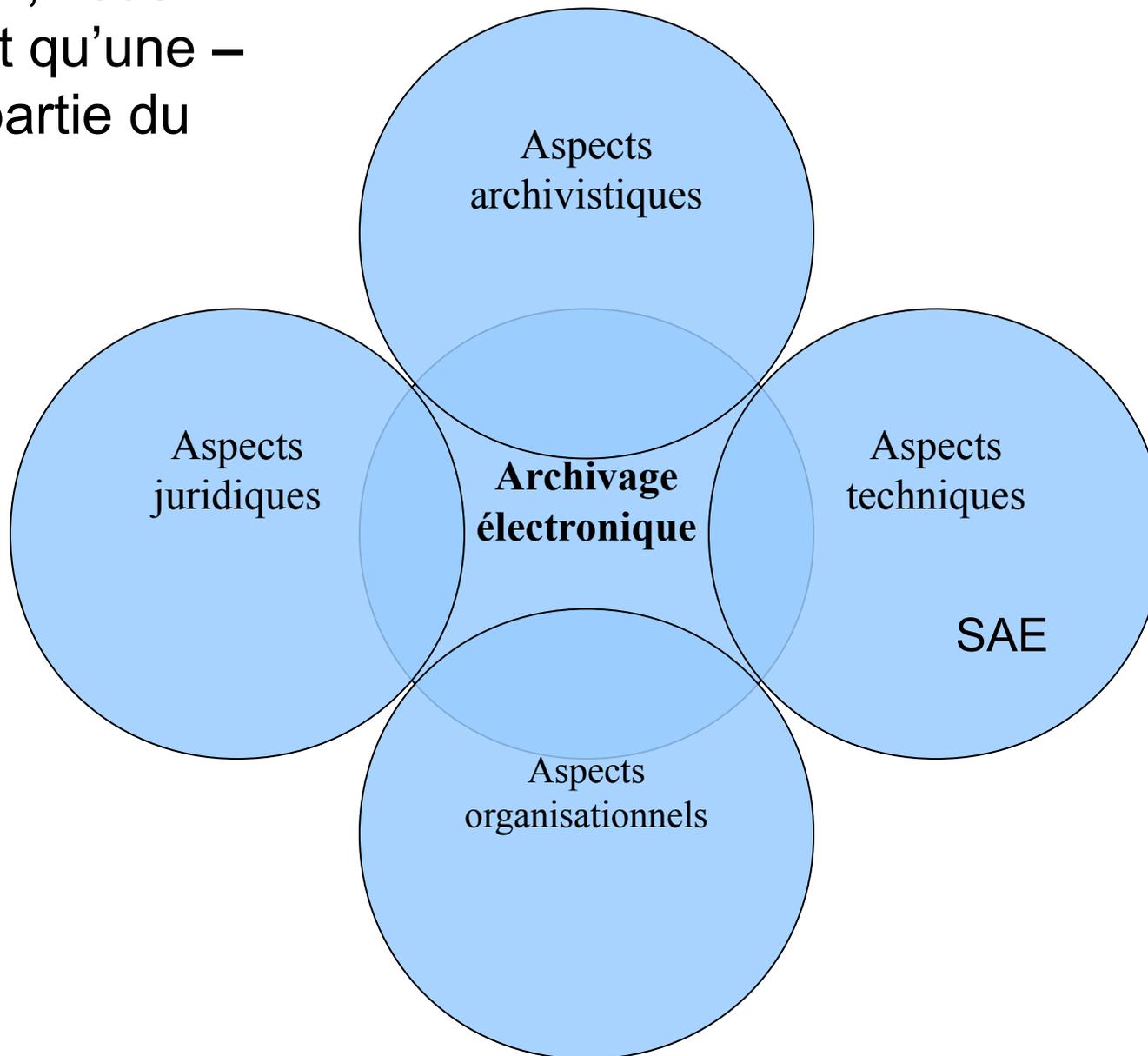


**Les supports de stockage amovibles sont conçus pour stocker et transporter des données à très court terme.**

Quand on produit des données numériques, il est indispensable de réfléchir à leur conservation et accessibilité afin de pouvoir les réutiliser à court, moyen et long terme.

Ce n'est pas seulement pour des raisons techniques

Si vous ne pensez  
qu'au SAE, vous  
n'avez fait qu'une –  
***petite*** – partie du  
travail ...



# *Le modèle OAIS*

## *(Open Archival Information System)*

- Standard élaboré dans l'univers aérospatial - Comité Consultatif pour les Systèmes de Données Spatiales (CCSDS)
- Né d'un réel besoin
- Norme ISO 14 721 (2003)
- Nouvelle version révisée en 2012

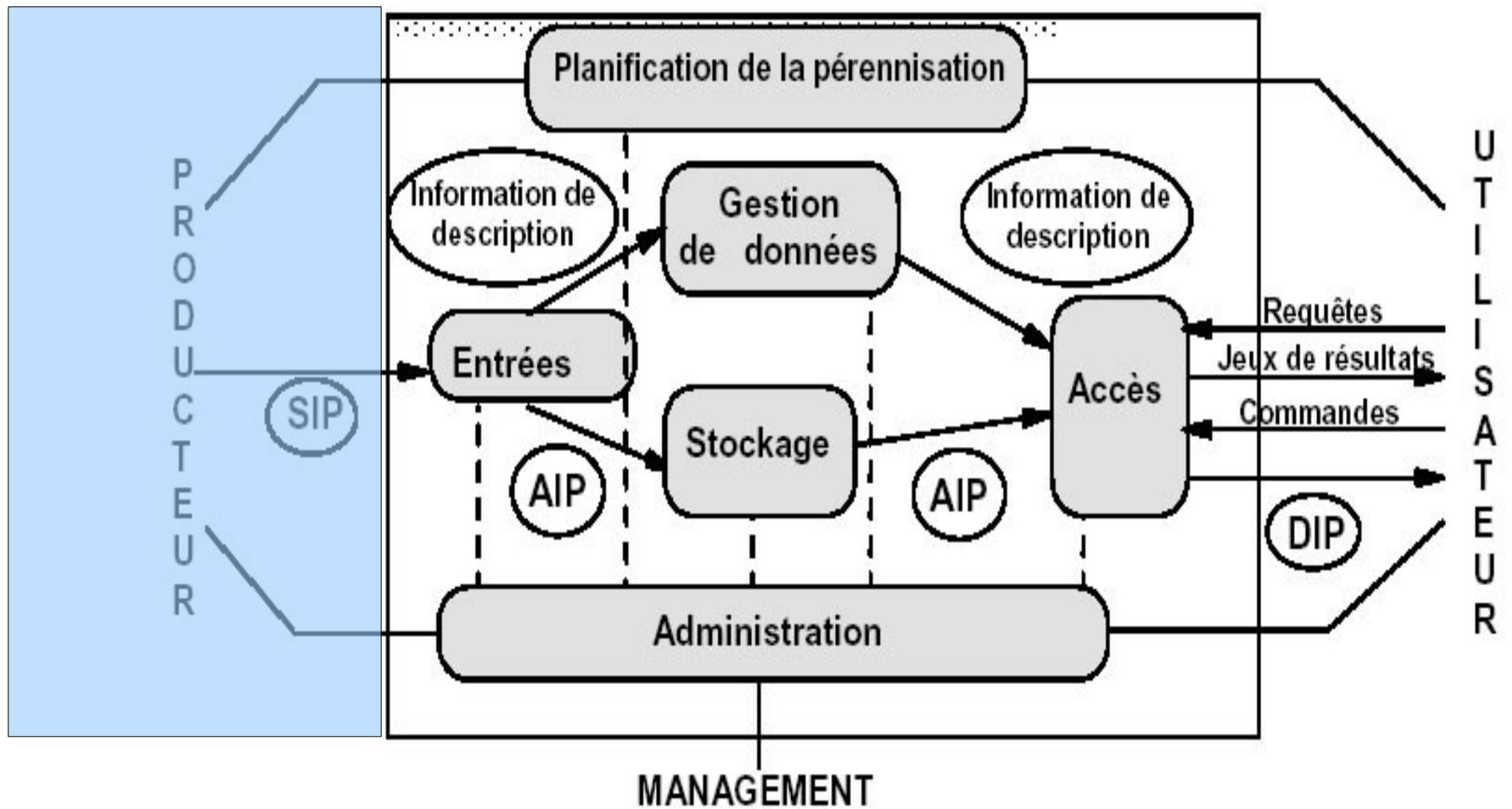


Schéma 4-1 : Entités fonctionnelles OAIS

# Trois éléments à prendre en considération:

1. Qualité des données produites

2. Cycle de vie de l'information  
(durées d'utilité, transferts de responsabilité)

3. Pérennisation

# *Le modèle OAIS*

## *Ses objectifs*

1. Comprendre les **enjeux et les moyens de la conservation numérique à long terme** et permettre de comparer différents systèmes
2. Définir des **concepts** et un vocabulaire & **modéliser un système** et l'information gérée par ce système

=> Mais pas de préconisations techniques ni de modèle clé en main

# *Avertissement*

« Le présent Modèle de référence s'applique à toutes les disciplines et organisations qui conservent et fournissent de l'information sous forme numérique (ou prévoient de le faire).

Par conséquent, les termes définis ci-dessous peuvent ne pas correspondre à tous les termes d'une discipline particulière (par exemple, les archives traditionnelles, les bibliothèques numériques, les centres de données scientifiques).

L'approche retenue est celle d'utiliser des termes non encore chargés de signification, afin de réduire le risque de véhiculer des acceptions involontaires. Il sera donc nécessaire pour toutes les disciplines et toutes les organisations de trouver la correspondance entre certains de leurs termes propres les plus courants, et les termes du présent Modèle de référence OAIS. »

*(OAIS, 1.7.2 Terminologie)*

# *Le modèle OAIS*

## *Concepts et vocabulaire*

4 notions

=> Caractérisation de l'information

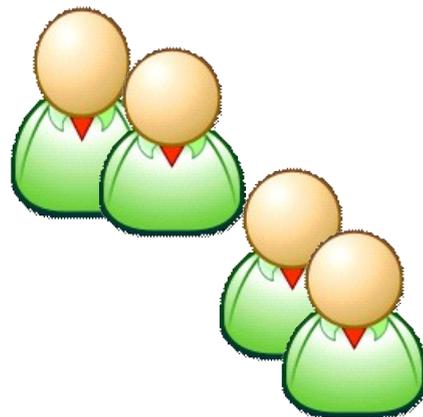


=> Paquets d'information



=> Archive

=> Acteurs



# *Caractérisation de l'information*



- Ce que l'on veut conserver : **l'information**
- forcément **codée** et **indépendante de son support**

=> ce que l'on conserve est donc la **représentation de l'information**

# *Caractérisation de l'information*



Mais caractériser l'information est plus complexe dans l'OAIS...

**Objet de données (OD) + Information de représentation (IR)**

= **Contenu d'informations** (minimum de ce qui peut être conservé et accessible sur le long terme)

# Caractérisation de l'information

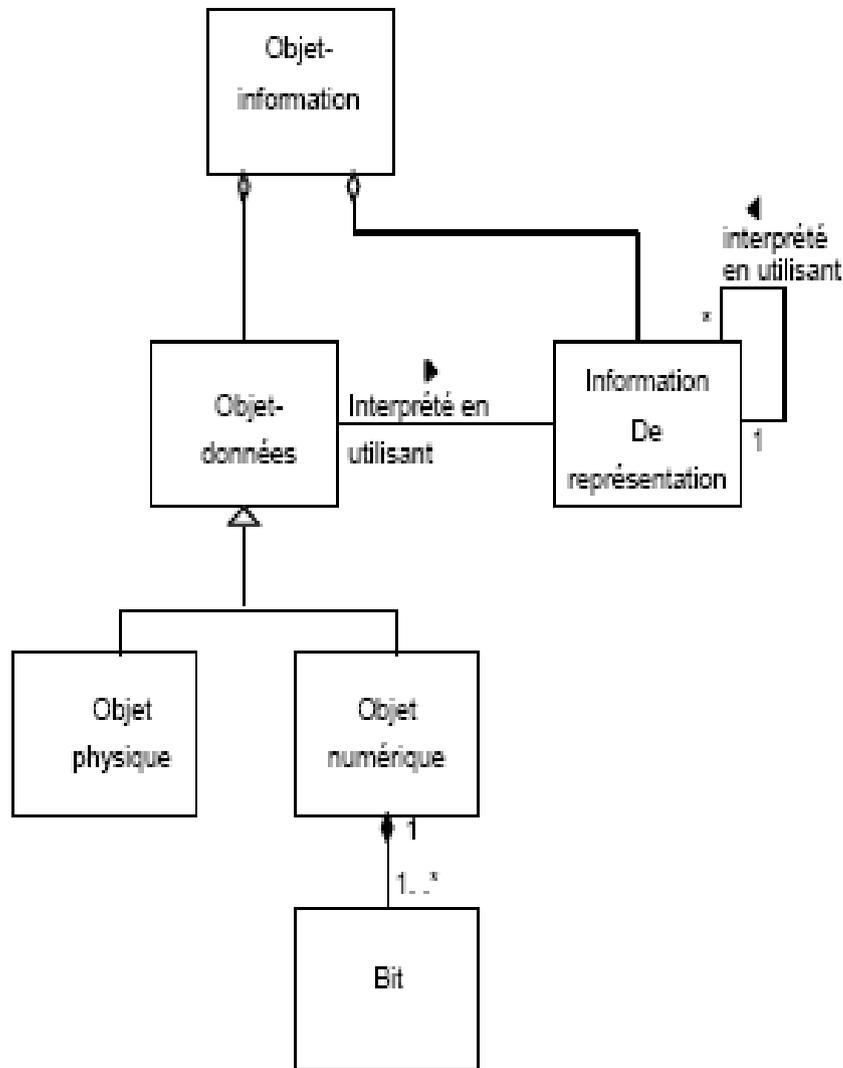


Schéma 4-10 : Objet-information

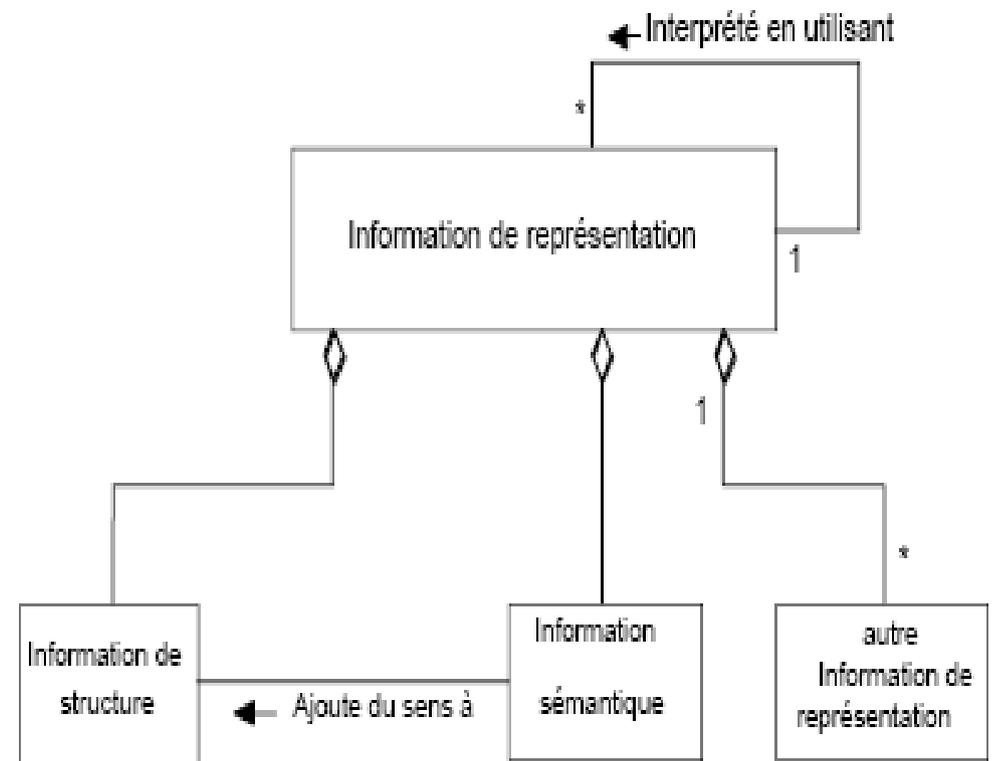


Schéma 4-11 : Objet-information de représentation

# Caractérisation de l'information



Exemples d'**information de représentation** :

- structure : format, version...; structure d'un fichier structuré
- exemple : numéro de sécurité sociale

<b>information</b>	<b>position de départ</b>	<b>longueur</b>	<b>position d'arrivée</b>
nsecsex	1	1	1
nsecannais	2	2	3
nsecmoinais	4	2	5
nsecdepnais	6	2	7
nseccommnais	8	3	10
nsecnordnais	11	3	13
nsecclle	14	2	15

- exemple : Instrument de recherche en XML/EAD  
(eadheader, archdesc, c...)

# Caractérisation de l'information



Exemples d'information de représentation :

- sémantique : langue(s) ; codes :

nsecsex	numéro de sécurité sociale, sexe
nsecannais	numéro de sécurité sociale, année de naissance
nsecmonais	numéro de sécurité sociale, mois de naissance
nsecdepnais	numéro de sécurité sociale, département de naissance
nseccommnais	numéro de sécurité sociale, commune de naissance
nsecnordnais	numéro de sécurité sociale, numéro d'ordre dans le mois de naissance
nseccl	numéro de sécurité sociale, clé de contrôle

- ex : dictionnaire des balises ead (unitid = cote, unittitle = intitulé...)

# *Caractérisation de l'information*



- et variables :

nsecsex	1 = homme ; 2 = femme
nsecannais	de 00 à 99
nsecmonais	de 01 à 12

- grâce à ces informations de représentation, il devient possible de comprendre les objets de données suivants :
  - 271027510202518
  - 153111302512536
  - 296011414800645

## *Caractérisation de l'information*



=> **L'information de représentation** permet de **déchiffrer l'objet de données.**

> **Où la retrouver ?** Dans des documents conservés par le producteur des données qui sont des **archives** (un ou plusieurs articles du versement).

=> Le contenu d'information : le plus petit élément archivable

Mais comment rattacher ces données à leur environnement de production ?

# *Caractérisation de l'information*



Pour y répondre : les **informations de pérennisation (PDI)**

1. Identification
2. Provenance
3. Contexte
4. Intégrité

NB : la PDI ressemble beaucoup à la description archivistique (mais pas seulement!)

Car

Paquet OAIS ≠ versement

# *Caractérisation de l'information*



On a donc :

- > Un **contenu d'informations (OD + IR)**
- > Ses **informations de pérennisation (PDI)**

Qui sont **ENCAPSULES** par une **information d'empaquetage**

Qui est elle-même accessible grâce à **une information de description**

# Caractérisation de l'information

Ainsi, on a le schéma OAIS suivant :

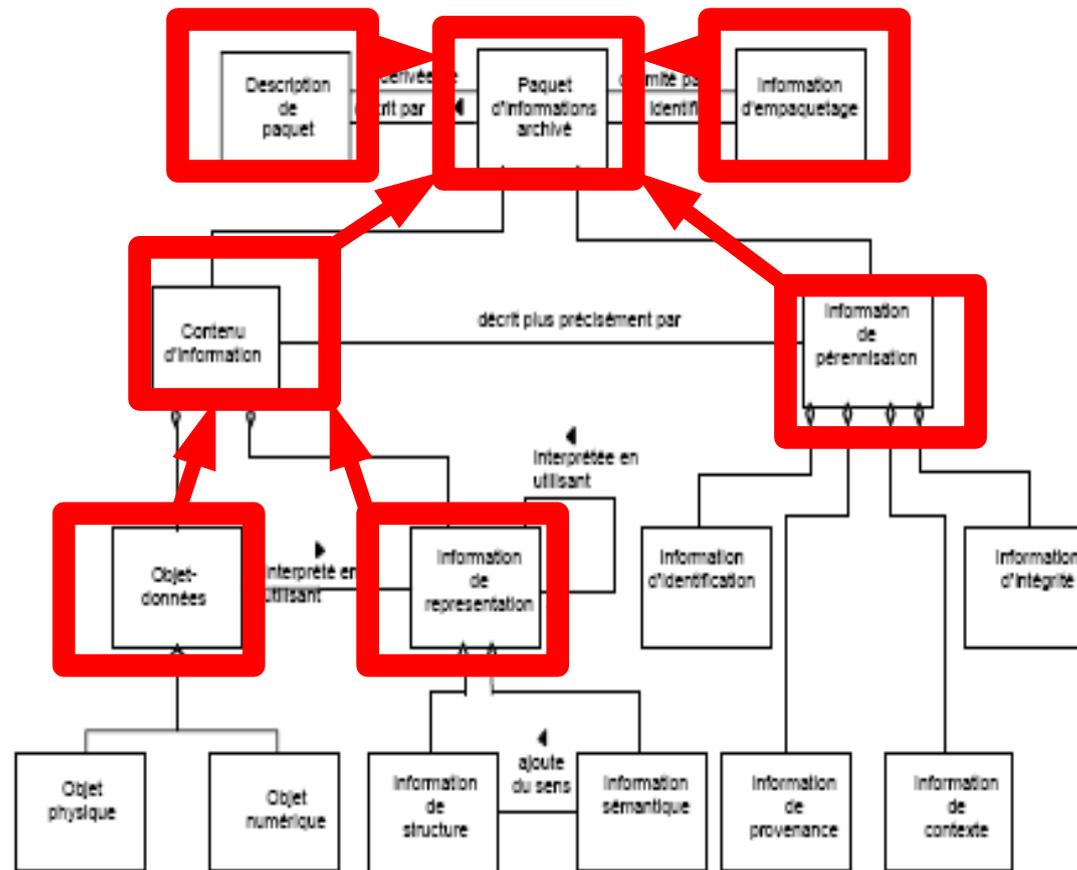


Schéma 4-18 : Paquet d'informations archivé (Vue détaillée)

# *Paquets d'information*



- Notion connexe :

Caractérisation des différents types de paquets d'informations selon 3 phases

- En entrée : **SIP**
- Conservation : **AIP**
- Diffusion : **DIP**

# Paquets d'information

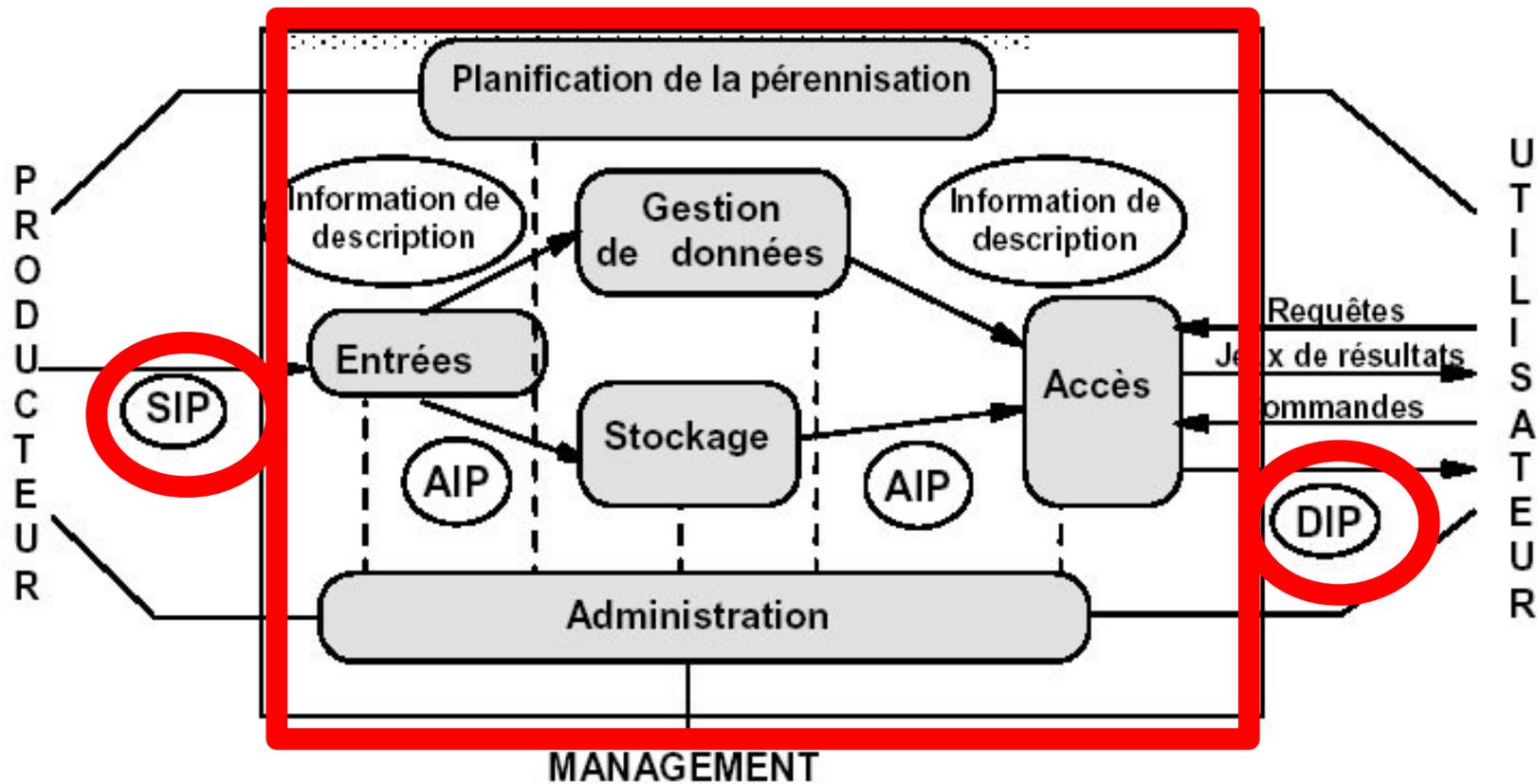


Schéma 4-1 : Entités fonctionnelles OAIS

## *Archive*

- Système mis en place pour conserver les données
- Process qui y sont associés
- Autorité responsable du système et des données

=> composé de **6 briques (entités OAIS)**

- Entrées
- Stockage
- Gestion des données descriptives
- Accès
- Administration
- Planification de la préservation



# Archive

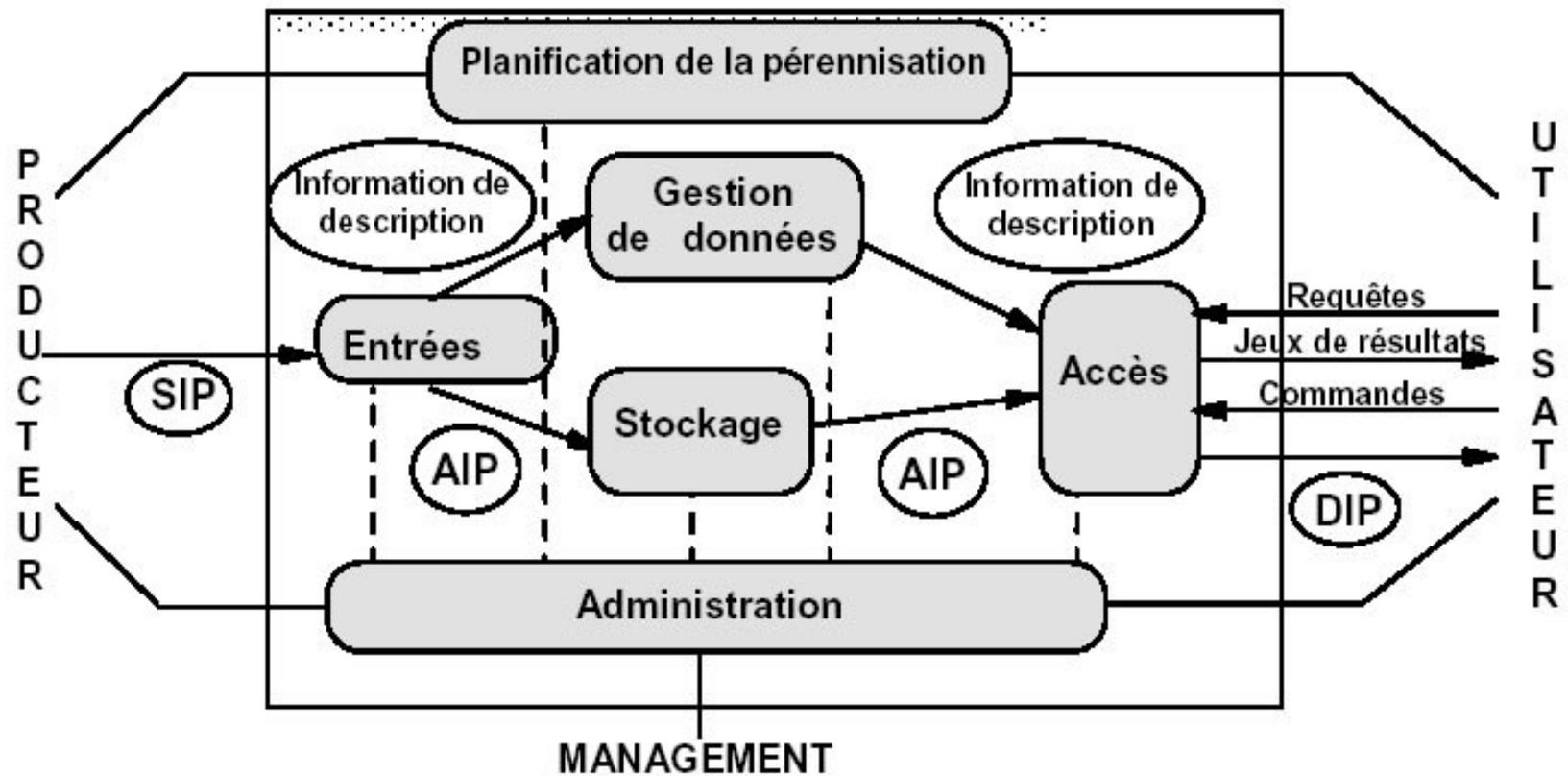


Schéma 4-1 : Entités fonctionnelles OAIS

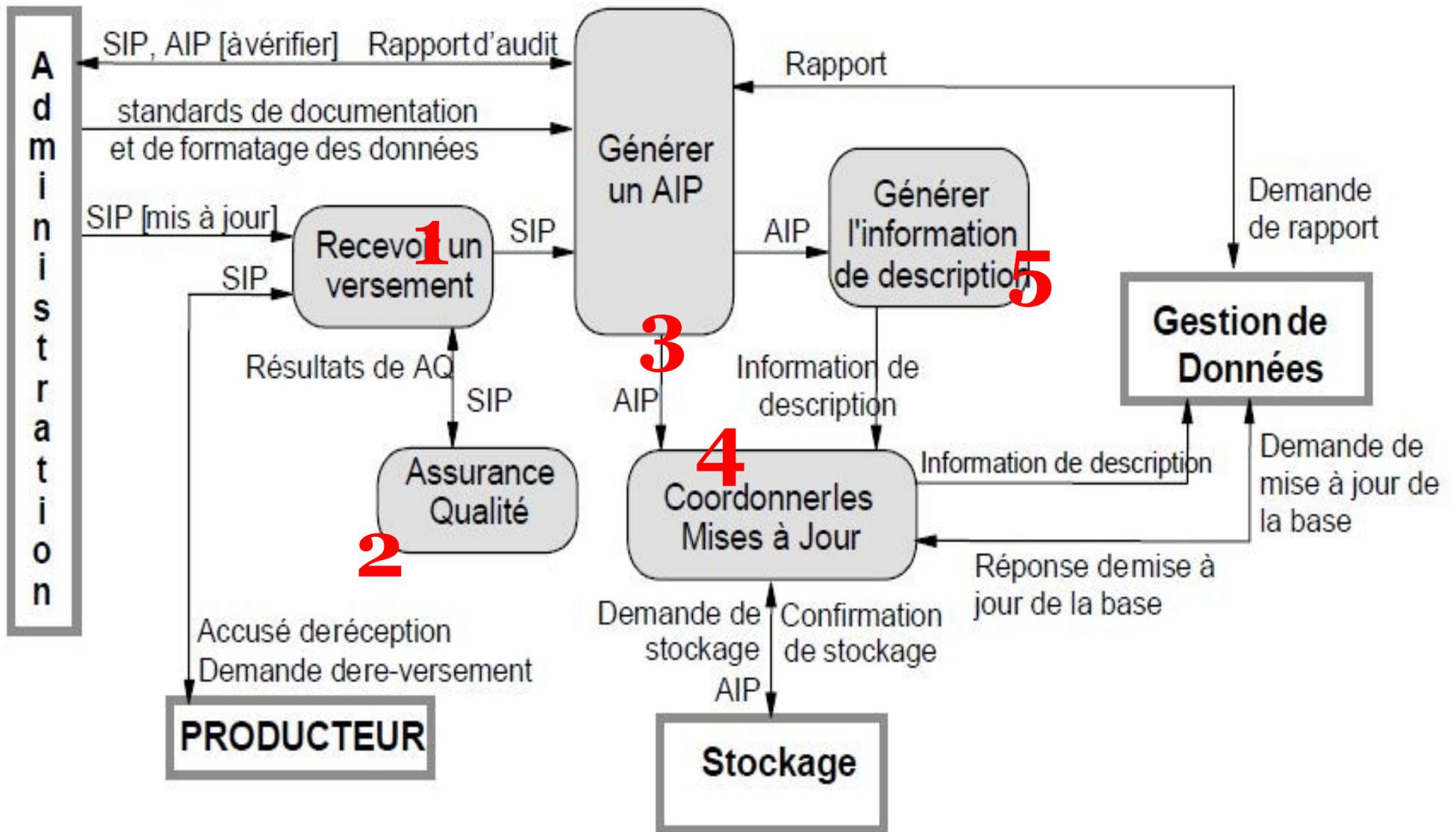
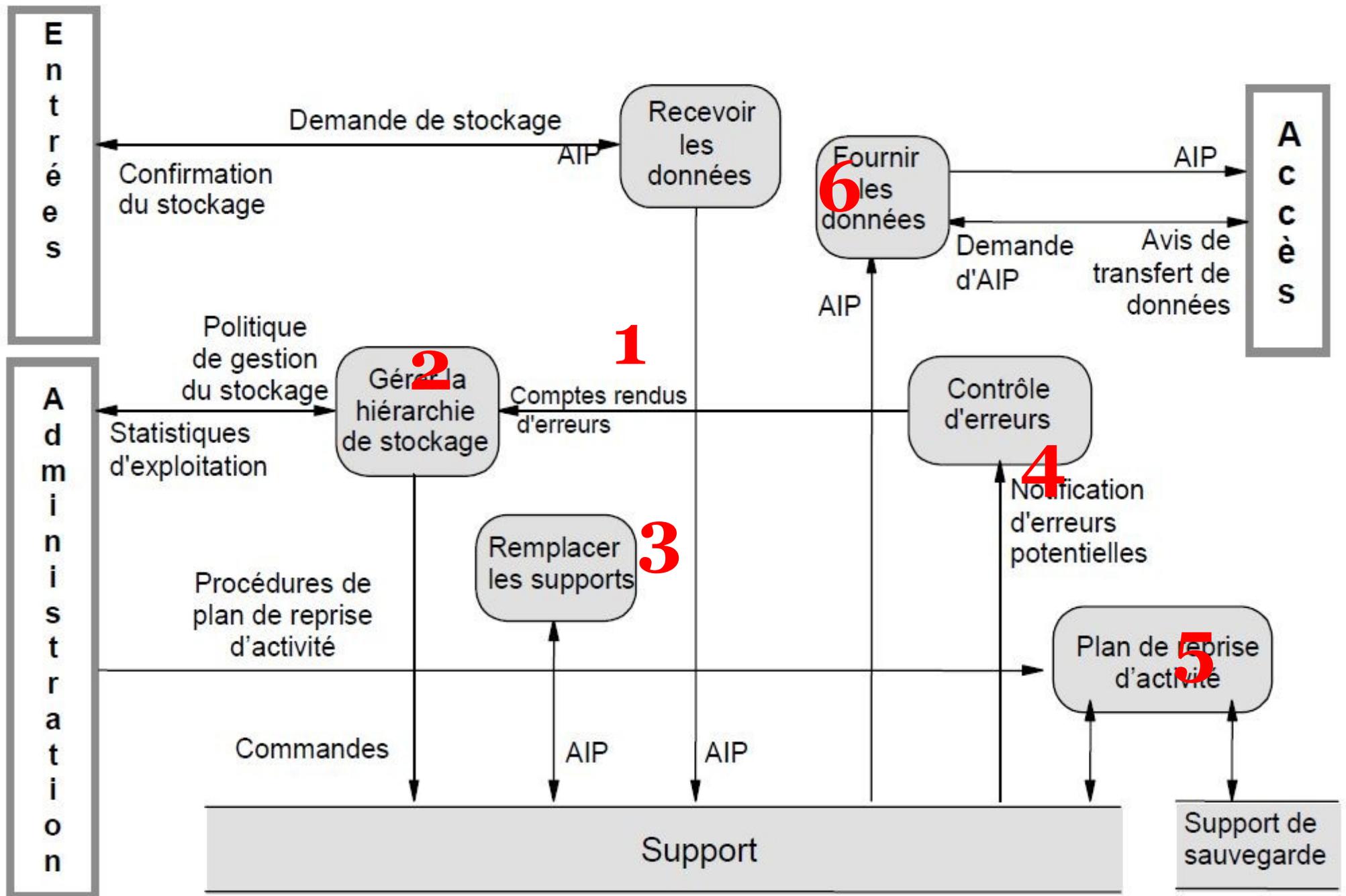
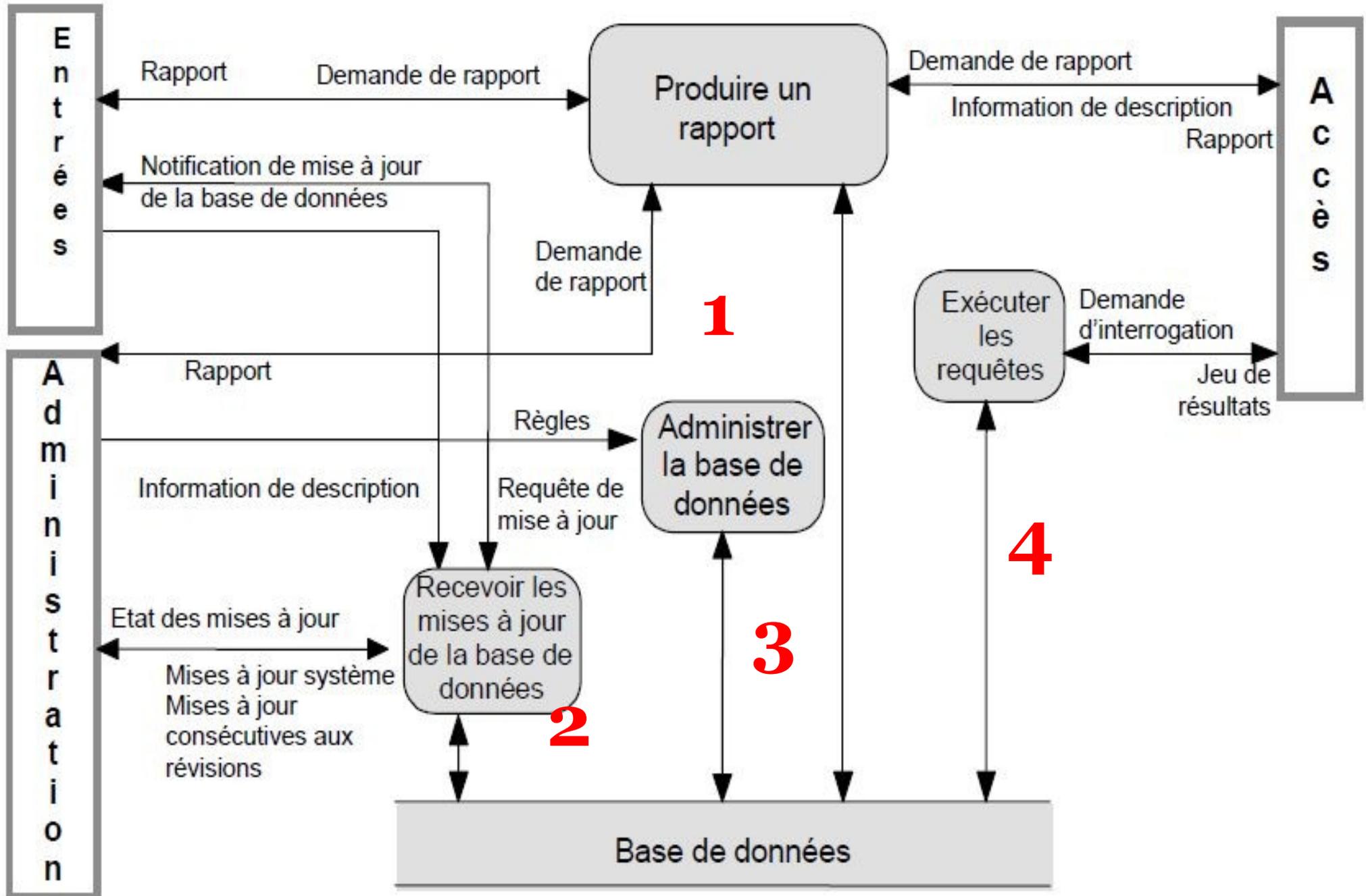


Schéma 4-2 : Fonctions de l'Entité « Entrées »



**Schéma 4-3 : Fonctions de l'Entité « Stockage »**



**Schéma 4-4 : Fonctions de l'Entité « Gestion de données »**

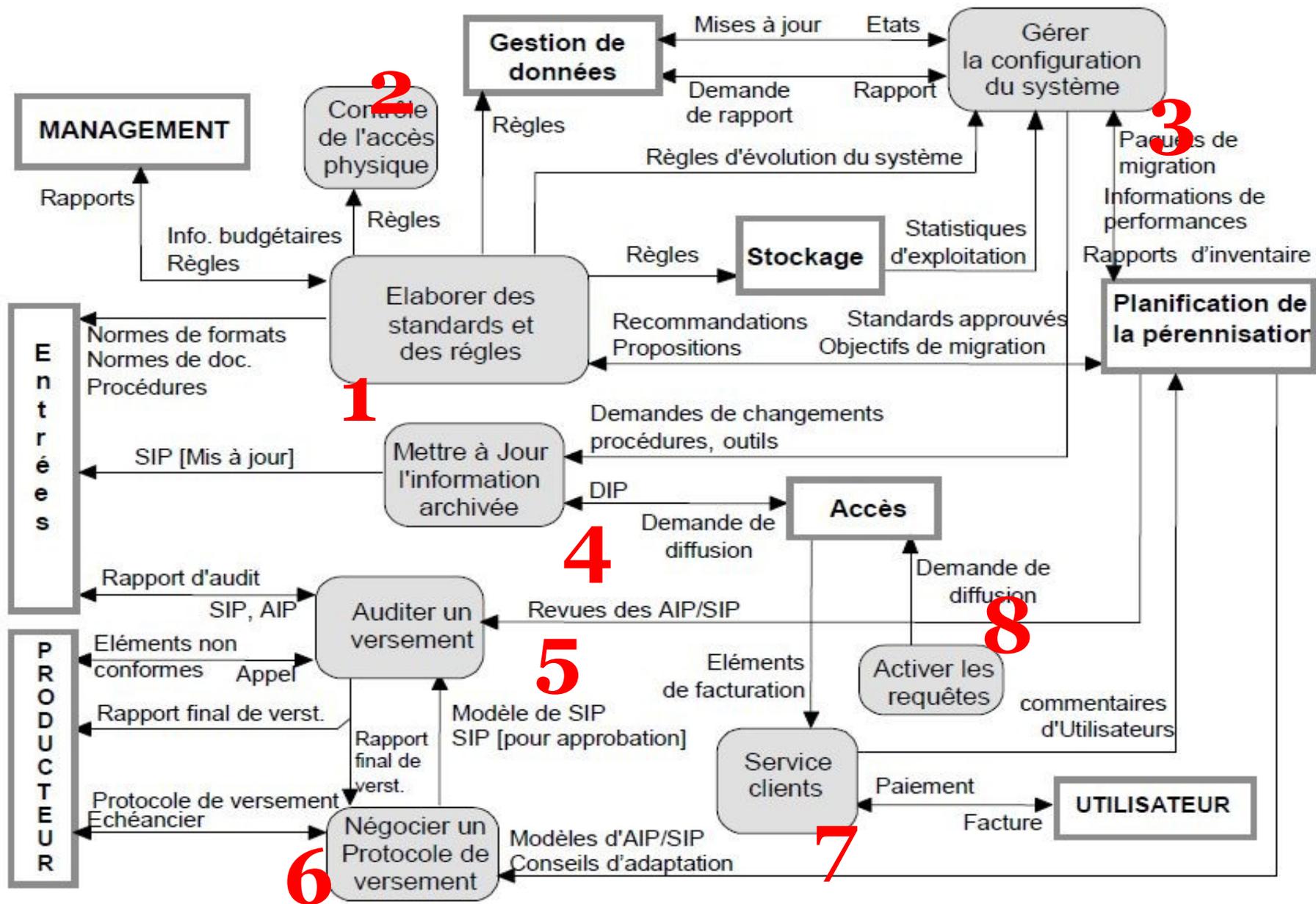


Schéma 4-5 : Fonctions de l'Entité « Administration »

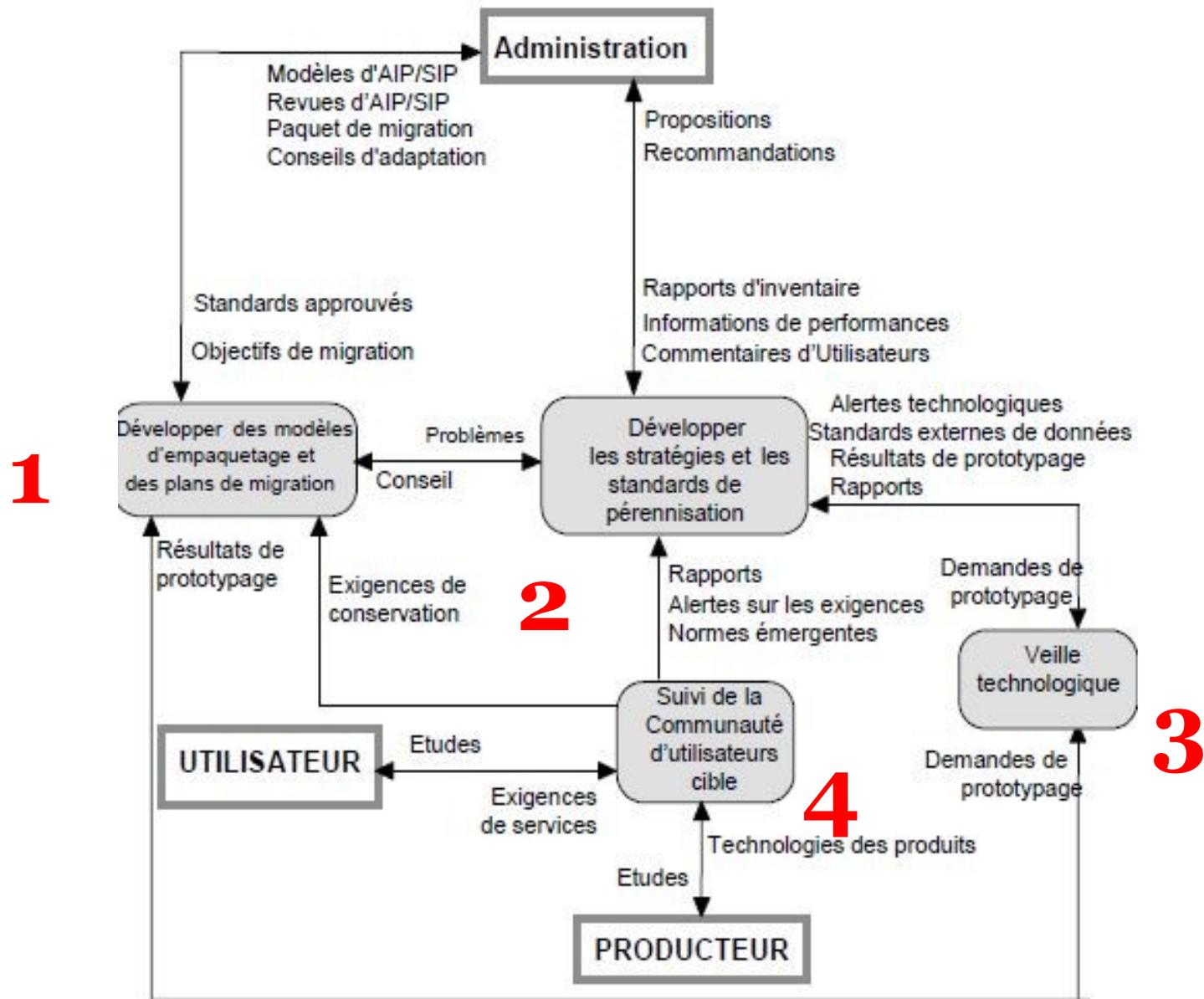


Schéma 4-6 : Fonctions de l'Entité « Planification de la pérennisation »

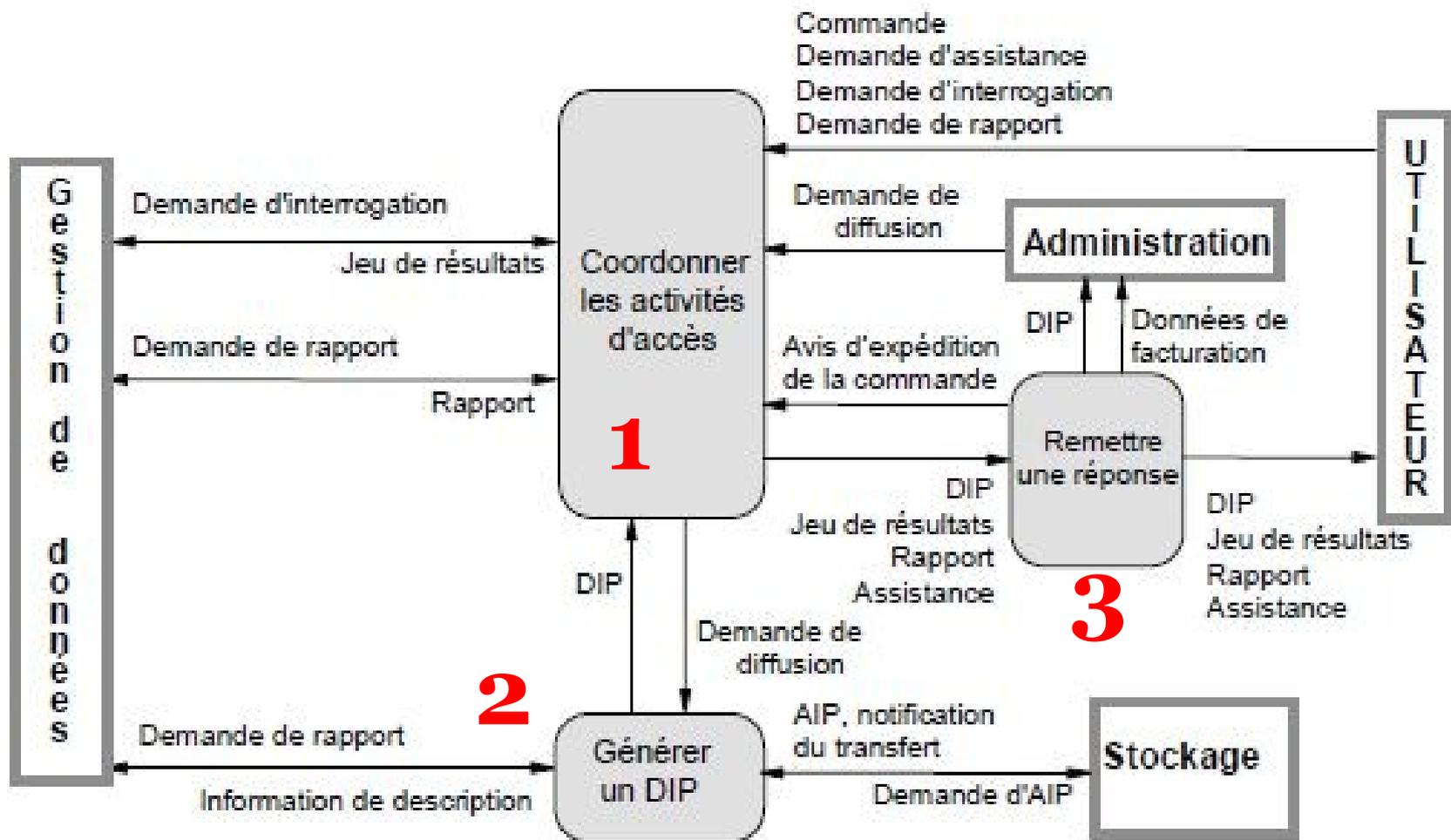


Schéma 4-7 : Fonctions de l'Entité « Accès »

# *Archive*

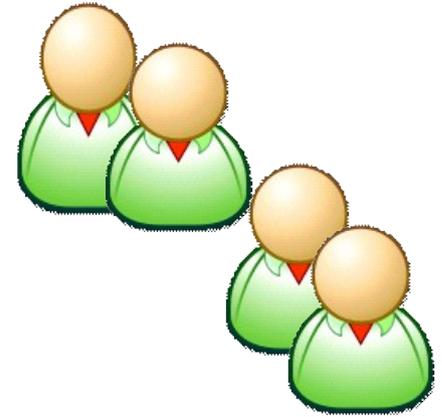


- Intérêt de ces 6 entités :

> Décomposition de la chaîne (par ailleurs très classique)

> Moyen d'analyser son propre système d'archivage

# *Acteurs*



- Producteur
- Management de l'Archive
- Communauté d'utilisateurs définie (base de connaissance commune à tous les utilisateurs)

# *Conclusions*

- Norme très utile mais caractère généraliste
- Un modèle à adapter à ses propres pratiques
- Un moyen d'analyser sa propre politique d'archivage électronique

Il faut donc mener une étude de ses besoins et de ses moyens et « **CONFIGURER** » l'**OAIS** selon le type de plateforme cible.

*NB : souvent citée dans les présentations des outils d'archivage électronique du marché*

## *Quelques recommandations:*

1. Identifier ses besoins et définir une stratégie
2. Délimiter un périmètre
3. Moduler les briques OAIS
4. Étudier l'impact des décisions prises sur les autres briques du système