



La collecte d'archives électroniques et Vitam



30/06/2022

Constats et enjeux



Le constat

Une **collecte complexe** des données :

- Non sécurisée (support amovible)
- Demandant de multiples interventions (supports informatiques)
 - processus chronophage et complexe
 - perte de la relation avec le producteur dans le contexte numérique

Des **SAE en production** qui contiennent **peu de données** à ce jour.

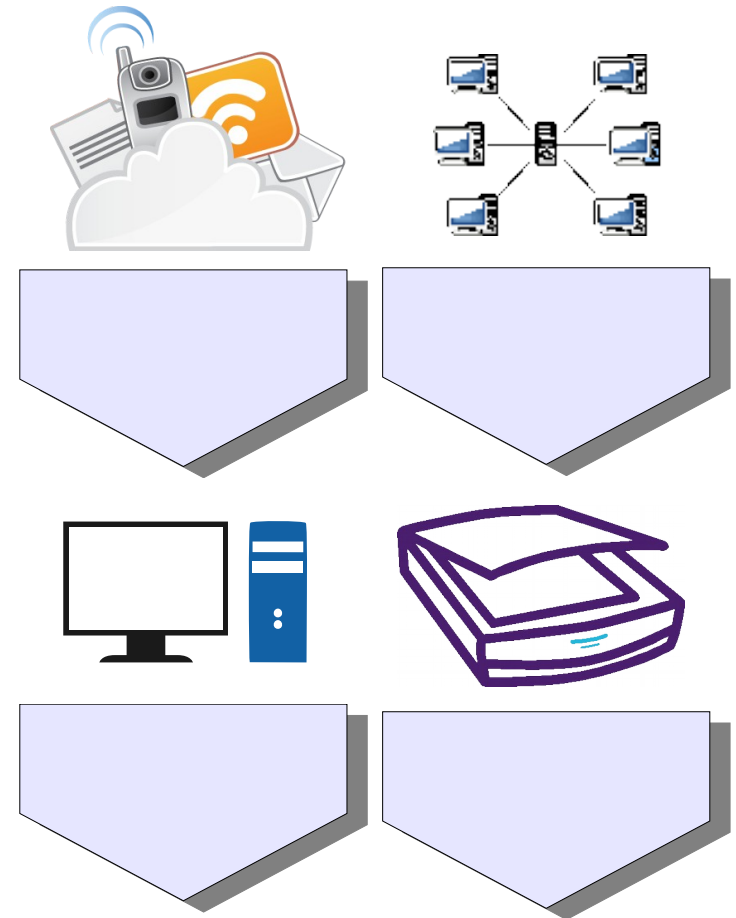
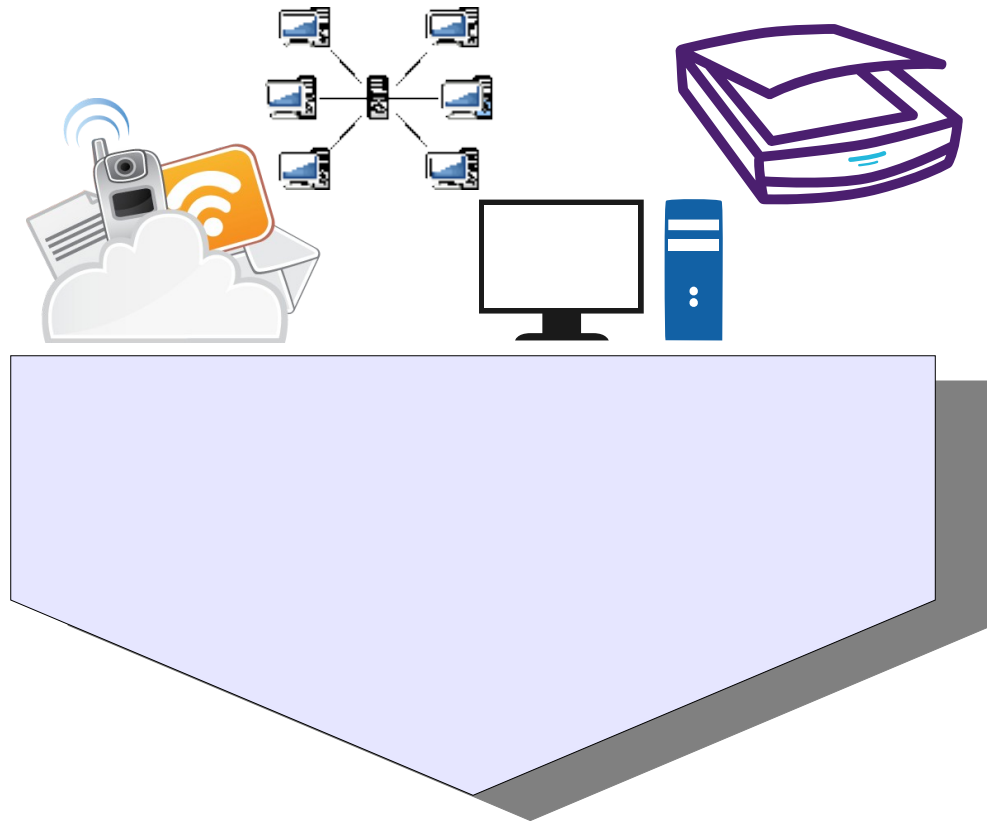
Des **serveurs** des administrations et des systèmes d'information contenant des données ayant vocation à être conservées dans des SAE

Nos enjeux : simplifier la collecte

- <ArchiveTransfer xsi:schemaLocation="fr:gouv:culture:archivesdefrance:seda.v2.0 seda-2.0-main.xsd" xml:id="ID1">
 - <Comment>Kit de prise en main</Comment>
 - <Date>2018-03-09T12:24:46</Date>
 - <MessageIdentifier>Kit de prise en main</MessageIdentifier>
 - <ArchivalAgreement>IC-000001</ArchivalAgreement>
- <CodeListVersions>
 - <ReplyCodeListVersion>ReplyCodeListVersion0</ReplyCodeListVersion>
 - <MessageDigestAlgorithmCodeListVersion>MessageDigestAlgorithmCodeListVersion0</MessageDigestAlgorithmCodeListVersion>
 - <MimeTypeCodeListVersion>MimeTypeCodeListVersion0</MimeTypeCodeListVersion>
 - <EncodingCodeListVersion>EncodingCodeListVersion0</EncodingCodeListVersion>
 - <FileFormatCodeListVersion>FileFormatC
 - <CompressionAlgorithmCodeListVersion>
 - <DataObjectVersionCodeListVersion>Data
 - <StorageRuleCodeListVersion>StorageRul
 - <AppraisalRuleCodeListVersion>Apprais
 - <AccessRuleCodeListVersion>AccessRuleC
 - <DisseminationRuleCodeListVersion>Diss
 - <ReuseRuleCodeListVersion>ReuseRuleCc
 - <ClassificationRuleCodeListVersion>Class
 - <AuthorizationReasonCodeListVersion>A
 - <RelationshipCodeListVersion>Relationshi
- <DataObjectPackage xml:id="ID2">
 - <BinaryDataObject id="ID5">
 - <DataObjectGroupId>ID4</DataObject

The screenshot displays a software interface for digital preservation. On the left, a tree view shows a hierarchy of files under 'Arbre du SIP (519 archivelnit/404 dog/404 hdo/0 pdo)'. The selected file is 'Adventures with ePub3_20180924-nn.pptx'. The main area shows a metadata editor for 'Unité d'archive - ID1490 ...' with a 'Niveau de description' table. A dropdown menu is open, listing various metadata fields such as 'Contenu textuel', 'Couverture', 'Date d'enregistrement', 'Date d'envoi', 'Date de création', 'Date de début', 'Date de fin', 'Date de numérisation', 'Date de réception', 'Date de transaction', 'Description', 'Destinataire', 'Destinataire-copie', 'Emetteur', 'Etat', 'Evénement', 'Expéditeur', 'GPS', 'Historique conservation', 'ID-ile-SI-prod', 'ID-SI-prod', 'ID-archiver', 'ID-classement-prod', 'ID-interne', 'ID-papier', 'ID-producteur', 'ID-versant', 'Langue', 'Langue de description', 'Mot-clé', 'Objet associé', 'Rédacteur', and 'Service producteur'. Buttons for 'Recharger unité d'archive' and 'Sauver unité d'archive' are visible.

Nos enjeux : faciliter l'interopérabilité





Au commencement,
le défi SEDAccord



Retour sur le défi SEDAccord

Proto Persona / Persona

UX RESEARCH - SEDACCORD



Patrick
Service
producteur
/ Versant

Age / Rôle

43 ans
Agent de l'administration
- Chargé de mission pour le projet "Trucmucho"

Mon histoire

Je crée de la donnée sur mon PC ou sur des espaces partagés mais sans en avoir conscience.

Je crée et reçoit des documents.

Je mets à jour des documents.

J'envoie des mails / Messages.

J'élimine quelques versoirs provisoires (parfois).

Parfois, ;

Parfois, ;

Quand j mes doi

Parfois, ;

Je m'ins



Sophie
Archiviste

Caractéristiques importantes

Informatiques
Connaissance du service d'archives
Connaissance des procédures **papiers**
Connaissance des procédures **électroniques**

Mes buts

Je ne veux pas passer beaucoup plus de temps pour verser, conserver et retrouver mes données.
Je suis preneur d'informations sur la gestion de mes documents.

Proto Persona / Persona

UX RESEARCH - SEDACCORD

Age / Rôle

30 ans
Archiviste

Mon histoire

Conseil aux services

Collecte d'archives électroniques et information sur le producteur

Création d'une entrée dans un SIA

Rattachement au plan de classement

Description du versement

Stockage : déplacement sur un espace de stockage intermédiaire

Recherche et communication

Élimination des données

Versement en tant qu'archives définitives

Caractéristiques importantes

Informatiques
Connaissance du service d'archives
Connaissance des procédures **papiers**
Connaissance des procédures **électroniques**

Mes buts

Je veux qu'on me verse des arborescences bureautiques contextualisées, que je puisse corriger et compléter et valider avant le versement d'un SAE.

Mes craintes

- Volumétrie des données
- Vrac
- Manque d'implication de services
- Multiplicité des outils
- Détails techniques et bidouillages

- Centré sur la collecte d'archives bureautiques
- 10 mois de défi pour une développeuse et un designer
- Phase d'analyse des pratiques et rédaction de persona
- Phase d'analyse des outils : ReSIP, Archifiltre, Gedeon, Octave, TSV
- Développement d'un POC

ARCHIVES COURANTES

ARCHIVES INTERMEDIAIRES

ARCHIVES DÉFINITIVES

CONSEIL

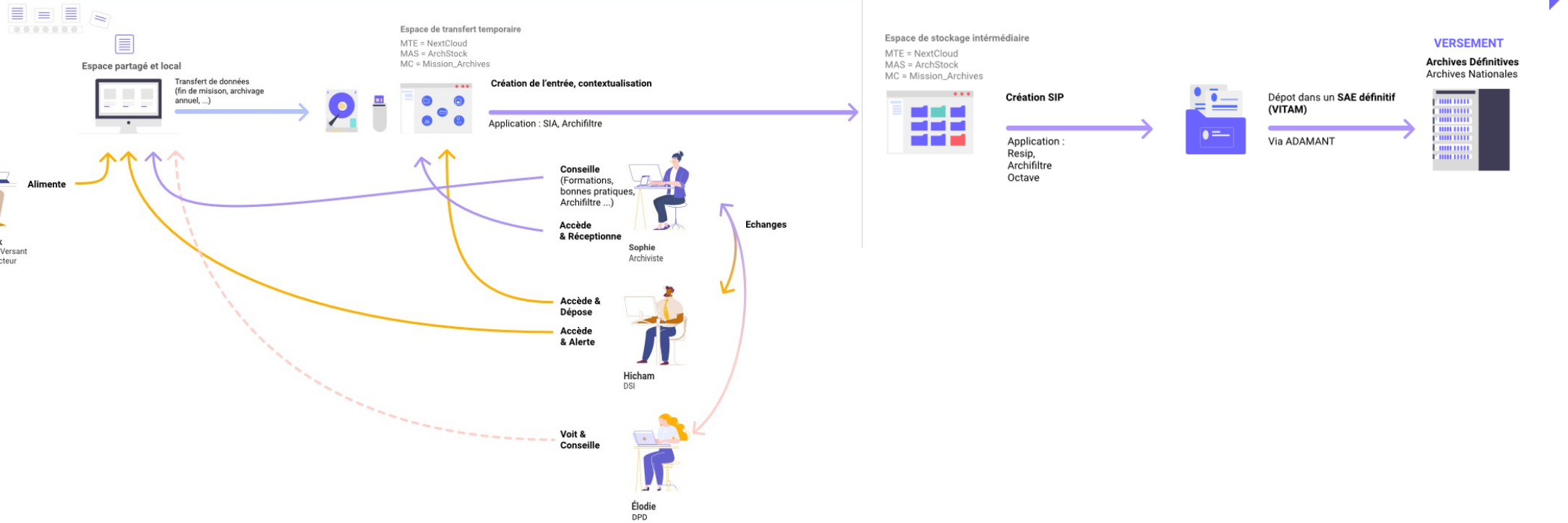
COLLECTE

TRAITEMENT

CONSERVATION INTERMEDIAIRE

TRAITEMENT, CREATION SIP

CONSERVATION DÉFINITIVE



Et après ?



Les suites du défi SEDAccord

Suite possible grâce à France Relance et au Service du numérique (SNum) du Ministère de la Culture

Industrialisation technique de SEDAccord

- Passer d'un démonstrateur sur le poste de travail à un outil en production
- Intégrer dans Vitam UI
- Clôturer la dette technique

Module de collecte complet

- SEDAccord
 - Automatiser des traitements techniques chronophages
 - Prendre en compte les versements volumineux
 - Intégrer de nouvelles fonctionnalités d'import
 - Prendre en compte les nouvelles fonctionnalités nécessaires aux utilisateurs
- Connecteur générique (en ré-utilisant le module SEDAccord et des travaux du CEA)
 - Développer un connecteur générique permettant de verser les données applicatives
 - Mettre en œuvre un flux depuis un SI prioritaire du Ministère de la Culture (SI-PAT)

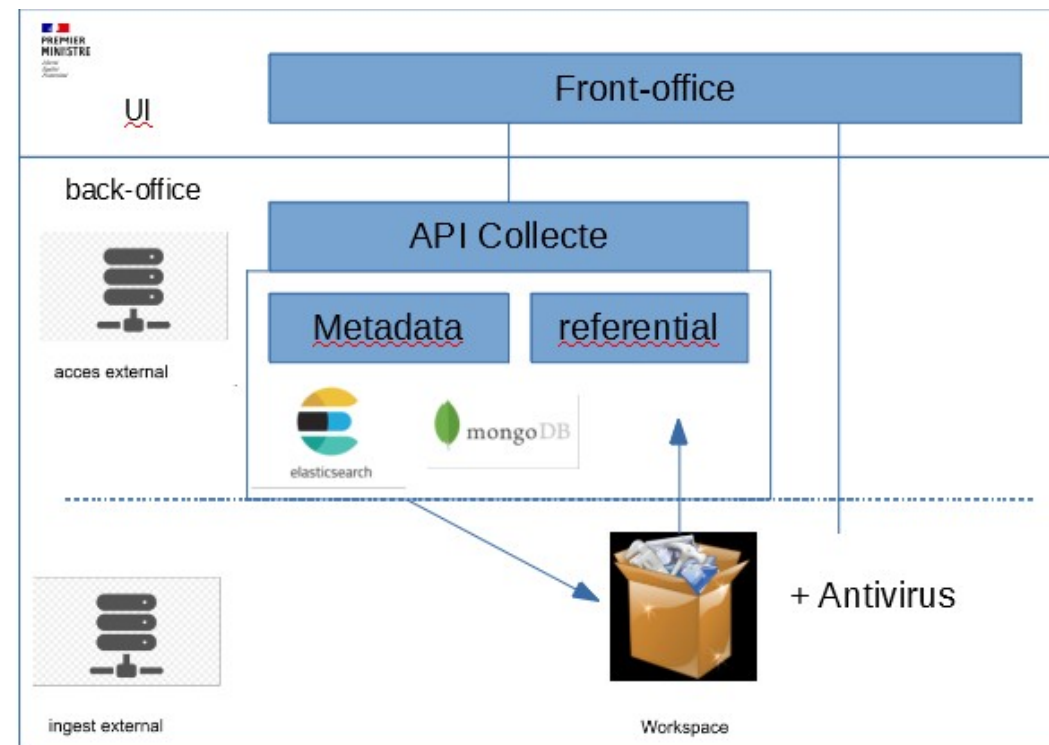
Élargissement du périmètre et des analyses

Des **échanges éclairant** sur les choix à opérer avec des retours d'expérience de la part du CEA, de l'INPI et du CINES/CNRS.

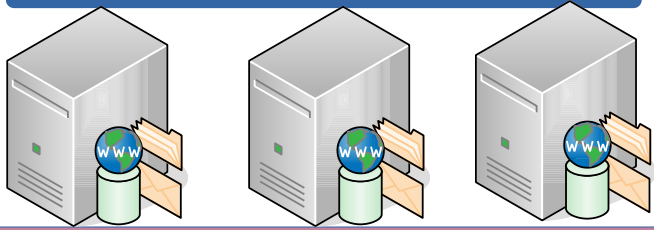
Phase de **conception technique** et d'architecture

Le module réutilise largement des composants internes afin de minimiser le coût de maintenance et permet d'envisager le développement de contrôles liés aux opérations d'entrée dans le module de collecte afin d'aider l'utilisateur à procéder à un versement conforme car déjà contrôlé.

Constitution d'un **backlog** d'évolutions souhaitées sur le périmètre du traitement des archives bureautiques (SEDAccord) et applicatives (connecteur générique) et partage avec l'équipe technique



VERSEMENT PAR FLUX



Système d'information de l'utilisateur

SIP

Données +
Fichiers

VITAM

Module de collecte

Base
Mongo

Espace de stockage

SIP

SIP

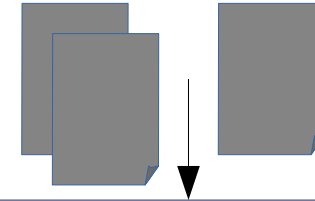
Modules de Vitam Core
Assurant la conservation
pérenne des archives

Base
Mongo

Espaces de stockage

VERSEMENT MANUEL

Arborescences
bureautiques



Autres types
de documents

VitamUI

APP Collecte

Données +
Fichiers

Module de collecte

Base
Mongo

Espace de stockage

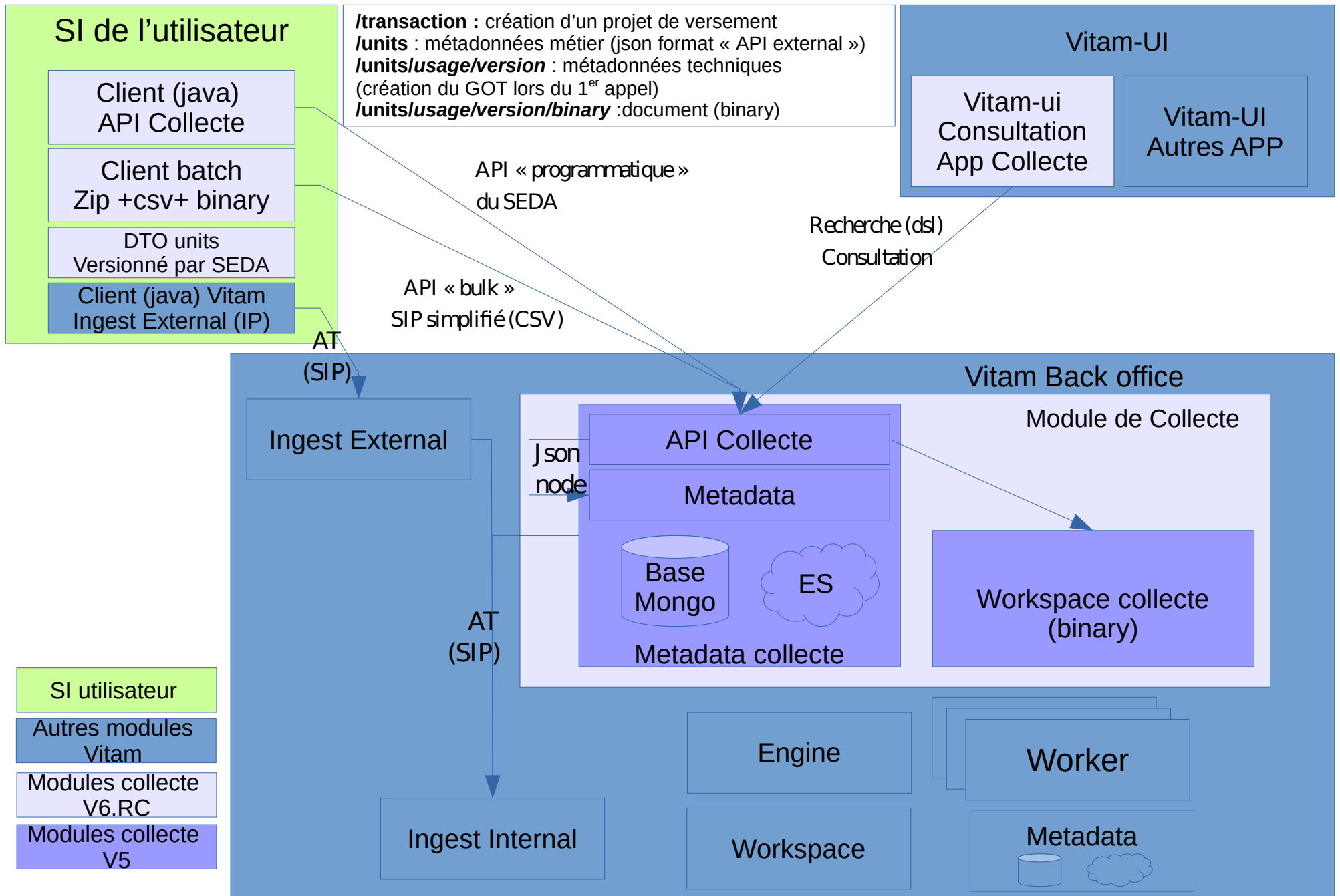
SIP

SIP

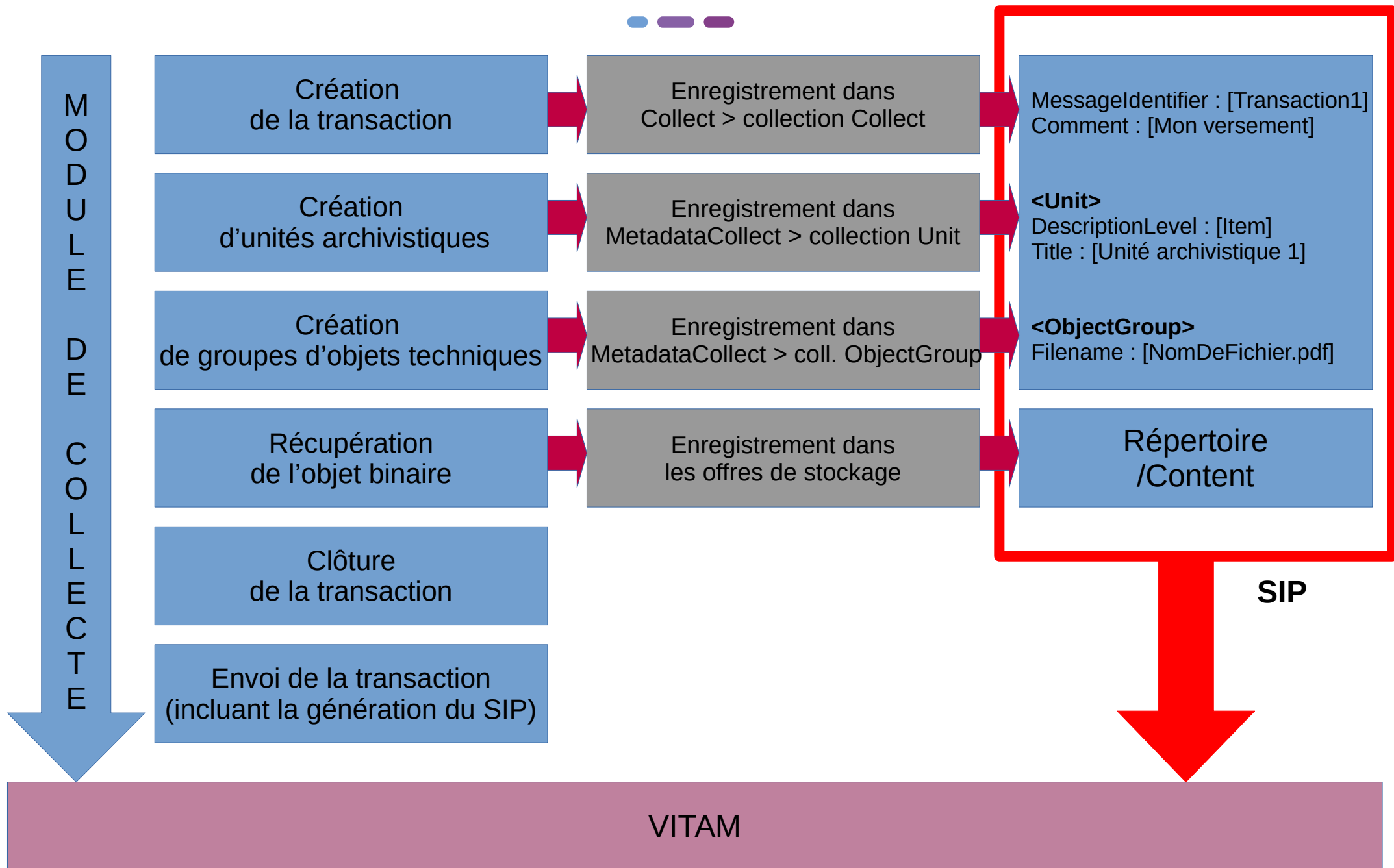
Modules de Vitam Core
Assurant la conservation
pérenne des archives

Base
Mongo

Espaces de stockage



Mécanisme mis en place



Démonstration d'un versement utilisant le module de collecte

En-tête

```
"ArchivalAgencyIdentifier": "Vitam",  
"TransferringAgencyIdentifier": "RATP",  
"OriginatingAgencyIdentifier": "RATP",  
"SubmissionAgencyIdentifier": "RATP",  
"MessageIdentifier": "20200131-000006",  
"ArchivalAgreement": "IC-000001",  
"Comment": "Bulletins de salaire : janvier 2020"
```

AU 1

```
"DescriptionLevel": "RecordGrp",  
"Title": "Bulletins de salaire : janvier 2020",  
"Tag": ["Paie", "Salaire", "Bulletin"],  
"StartDate": "2020-01-01",  
"EndDate": "2020-01-31",  
"#management": {  
  "AppraisalRule": {"Rules": [{"Rule": "APP-00001", "StartDate": "2020-01-31"}]},  
  "FinalAction": "Destroy",  
  "AccessRule": {"Rules": [{"Rule": "ACC-00001", "StartDate": "2020-01-31"}]}  
}
```

AU 2

```
"DescriptionLevel": "Item",  
"Title": "Dupont, Charles",  
"OriginatingAgencyArchiveUnitIdentifier": ["BP_RATP103058_20200128"],  
"TransactedDate": "2020-01-28",  
"#unitups": ["aeaaaaaaaghefnffaaxa4ama6vfwaeqaaaaq"],  
"#management": {  
}
```

GOT

```
"fileInfo": {  
  "filename": "BP_RATP103058_20200128.pdf"  
}
```

OT

ID909.pdf



Et l'archiviste dans
tout cela ?



Faciliter les aspects techniques et se recentrer sur les aspects fonctionnels

Processus de connexion d'un flux applicatif

Initialisation

Cadrage

Réalisation

Mise en fonction

Utilisation

Retrait de service ou procédure de reprise de données

Programme Vitam

47 rue de la Chapelle, 75018 Paris – France

Tél. : +33 (0)1 86 69 60 25

www.programmevitam.fr

<https://twitter.com/@ProgVitam>

<https://www.linkedin.com/grps/Programme-Vitam>

