Quand Vitam s’engage pour faciliter la collecte d’archives électroniques

25/05/2022

Éléments de contexte

Le constat

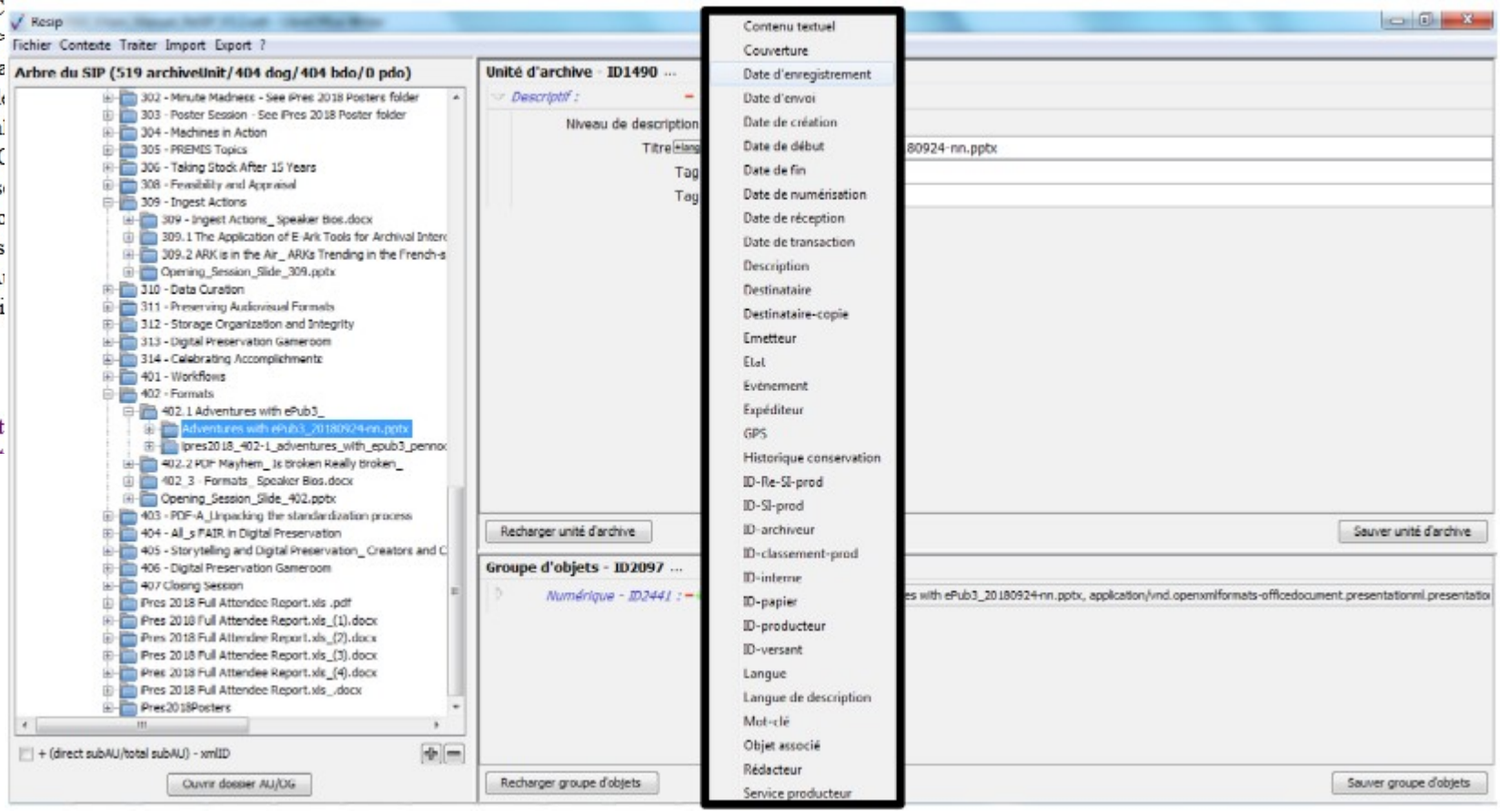
Une **collecte complexe** des données :
- Non sécurisée (support amovible)
- Demandant de multiples interventions (supports informatiques)
→ processus chronophage et complexe
→ perte de la relation avec le producteur dans le contexte numérique

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

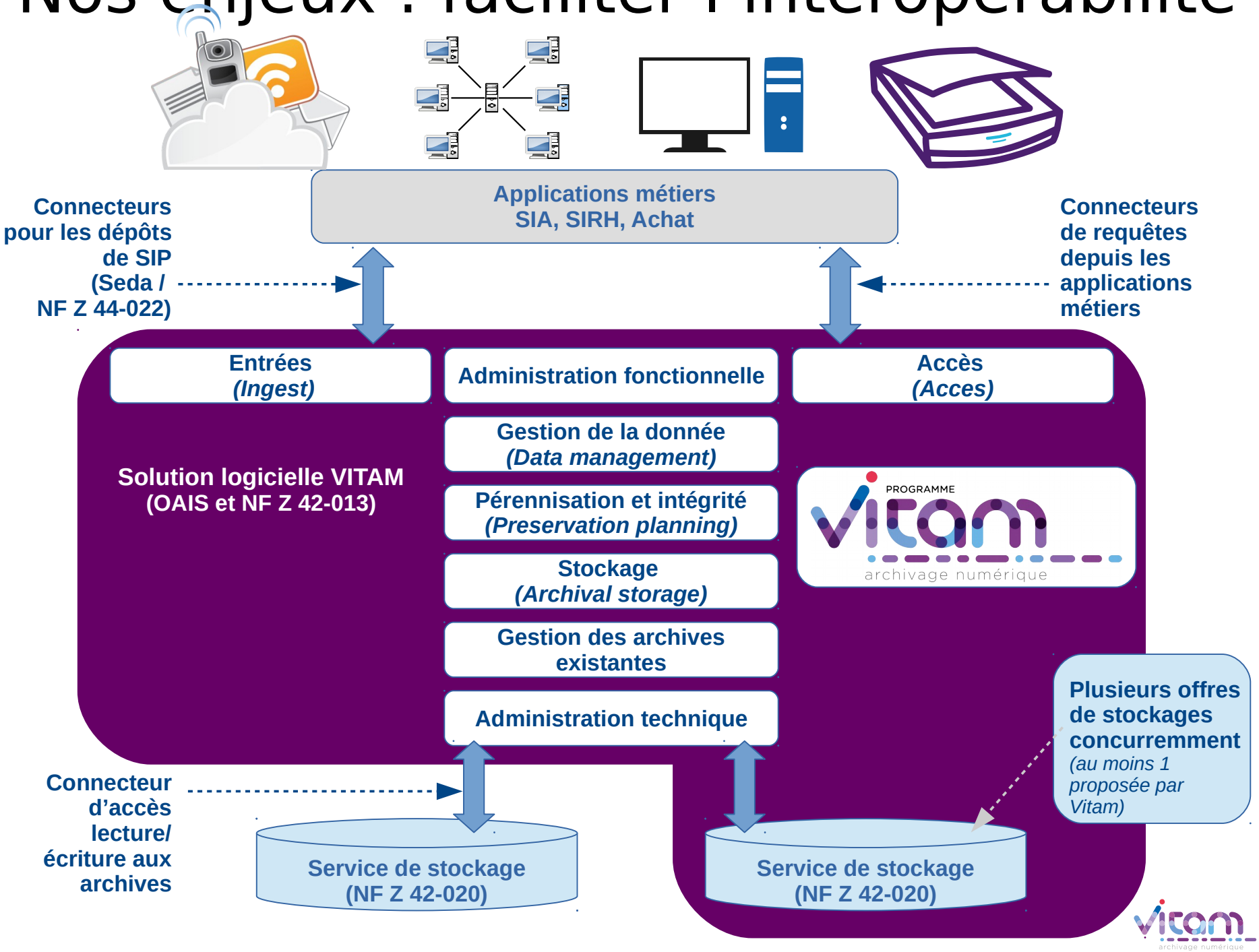
Des **serveurs** des administrations et des systèmes d'information contenant des données ayant vocation à être conservées dans des SAE

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>Appraisa
      <AccessRuleCodeListVersion>AccessRuleC
      <DisseminationRuleCodeListVersion>Diss
      <ReuseRuleCodeListVersion>ReuseRuleCc
      <ClassificationRuleCodeListVersion>Class
      <AuthorizationReasonCodeListVersion>A
      <RelationshipCodeListVersion>Relationshi
    </CodeListVersions>
  – <DataObjectPackage xml:id="ID2">
    – <BinaryDataObject id="ID5">
        <DataObjectGroupId>ID4</DataObject
```

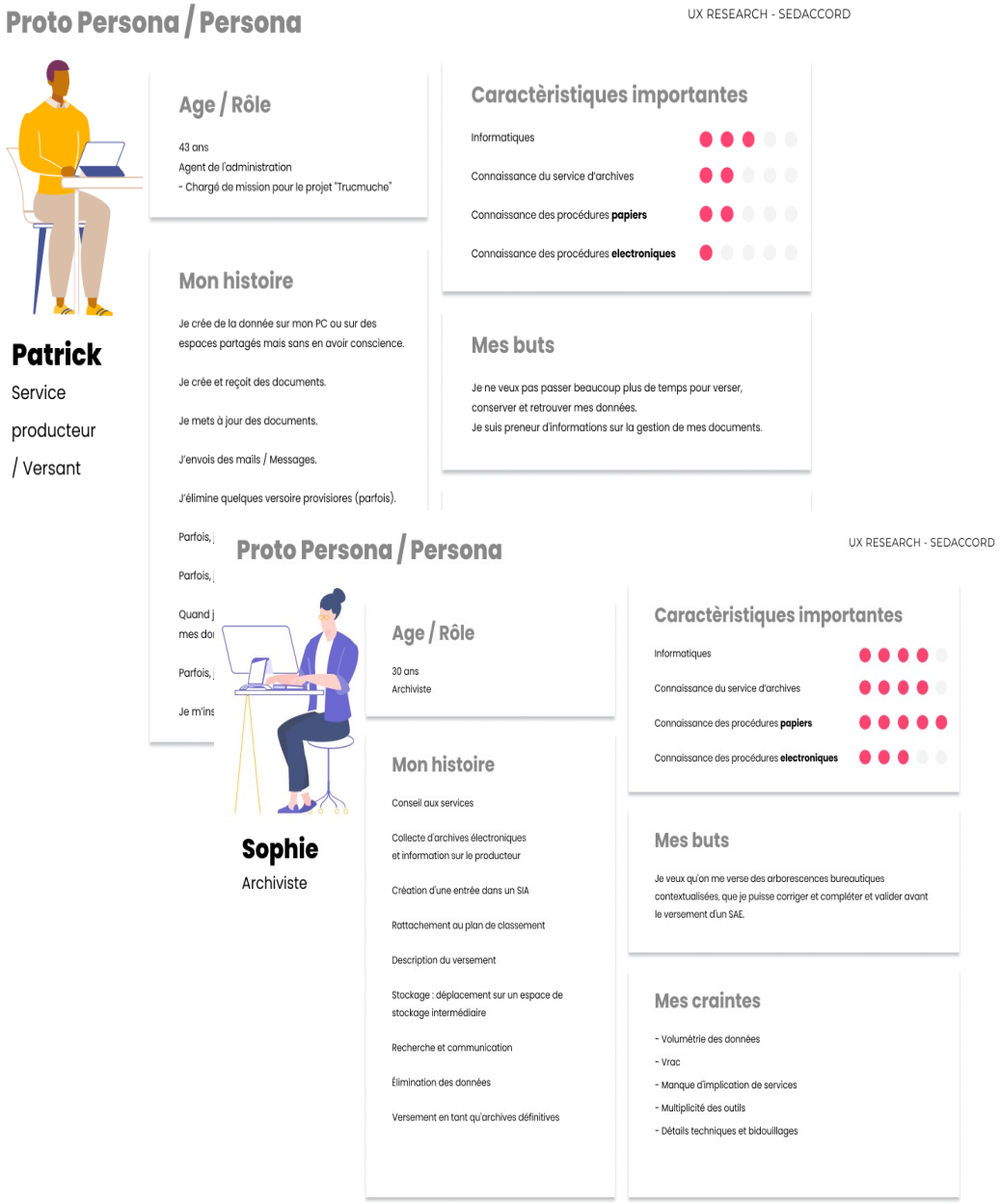



Nos enjeux : faciliter l'interopérabilité



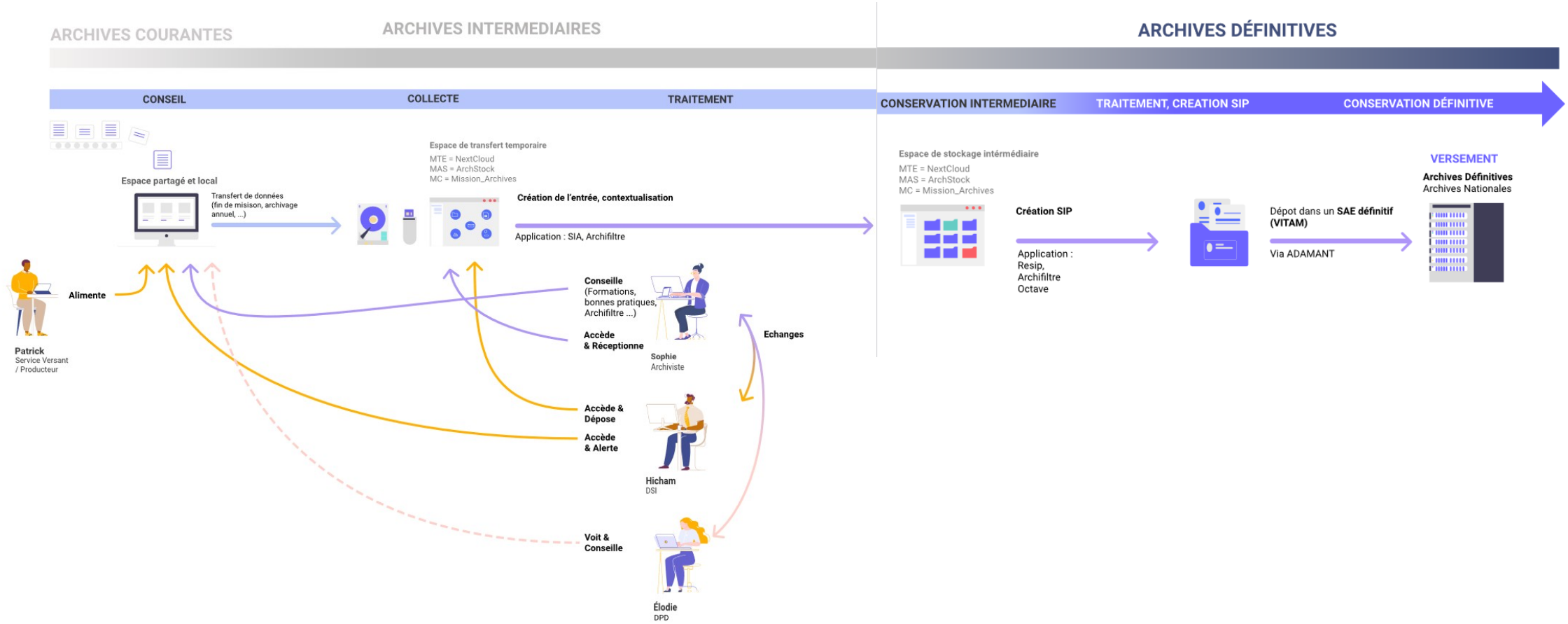
Au commencement, le défi SEDAccord

Retour sur le défi SEDAccord



- 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

vitam
archivage numérique


ARCHIVES COURANTES
ARCHIVES INTERMEDIAIRES
ARCHIVES DÉFINITIVES
CONSEIL
COLLECTE
TRAITEMENT
CONSERVATION INTERMEDIAIRE
TRAITEMENT, CREATION SIP
CONSERVATION DÉFINITIVE
Espace partagé et local
Transfert de données
(fin de misison, archivage
annuel, ...)
Espace de transfert temporaire
MTE = NextCloud
MAS = ArchStock
MC = Mission_Archives
Création de l'entrée, contextualisation
Application : SIA, Archifiltre
Espace de stockage intérmédiaire
MTE = NextCloud
MAS = ArchStock
MC = Mission_Archives
Création SIP
Application :
Resip,
Archifiltre
Octave
Dépot dans un SAE définitif
(VITAM)
Via ADAMANT
VERSEMENT
Archives Définitives
Archives Nationales
Alimente
Patrick
Service Versant
/ Producteur
Conseille
(Formations,
bonnes pratiques,
Archifiltre ...)
Accède
& Réceptionne
Sophie
Archiviste
Echanges
Accède &
Dépose
Accède
& Alerte
Hicham
DSI
Voit &
Conseille
Élodie
DPD

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

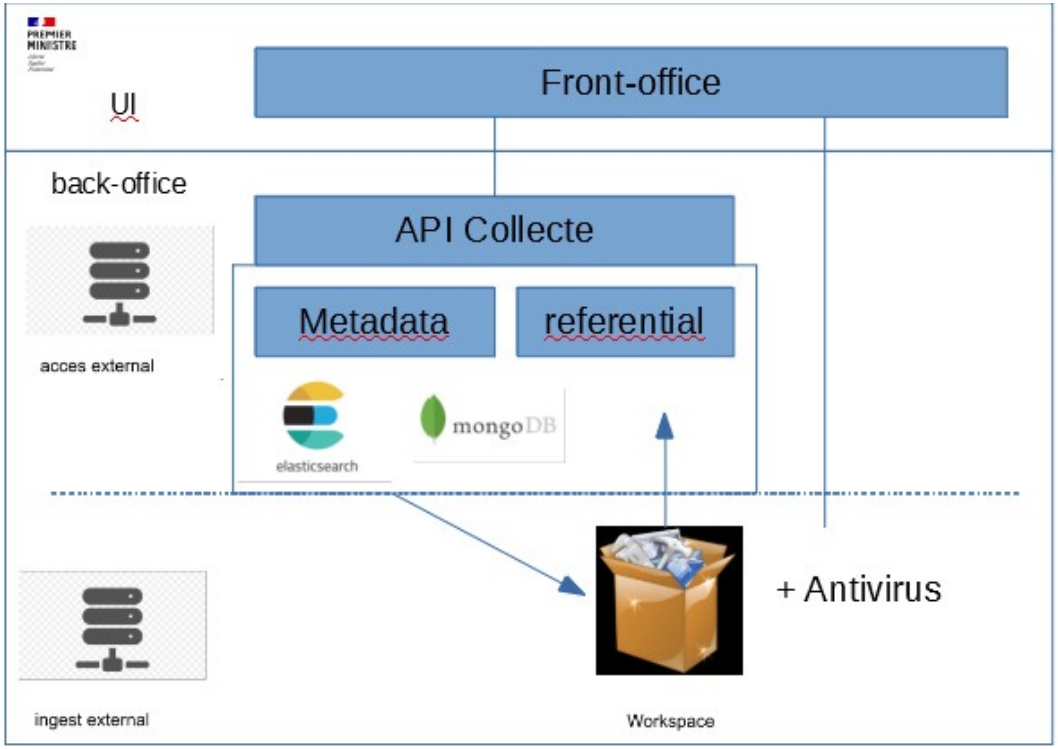
- *SEDAccord*
  - → Automatiser des traitements techniques chronophages
  - → Prendre en compte les versements volumineux
  - → Intégrer de nouvelles fonctionnalités d'import
  - → Prendre en compte les nouvelles fonctionnalités nécessaires aux utilisateurs
- *Connecteur générique (en ré-utilisant le 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

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

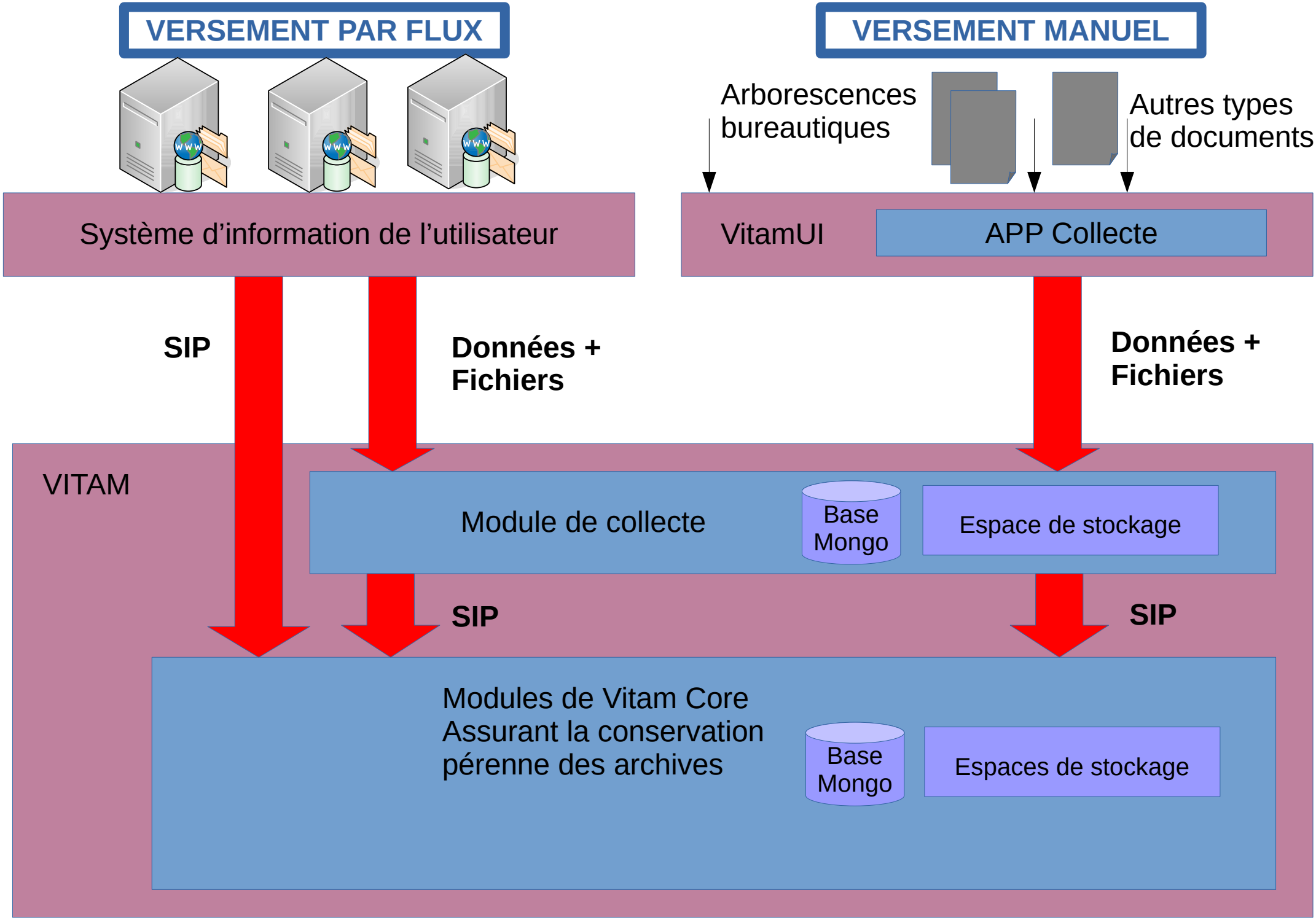


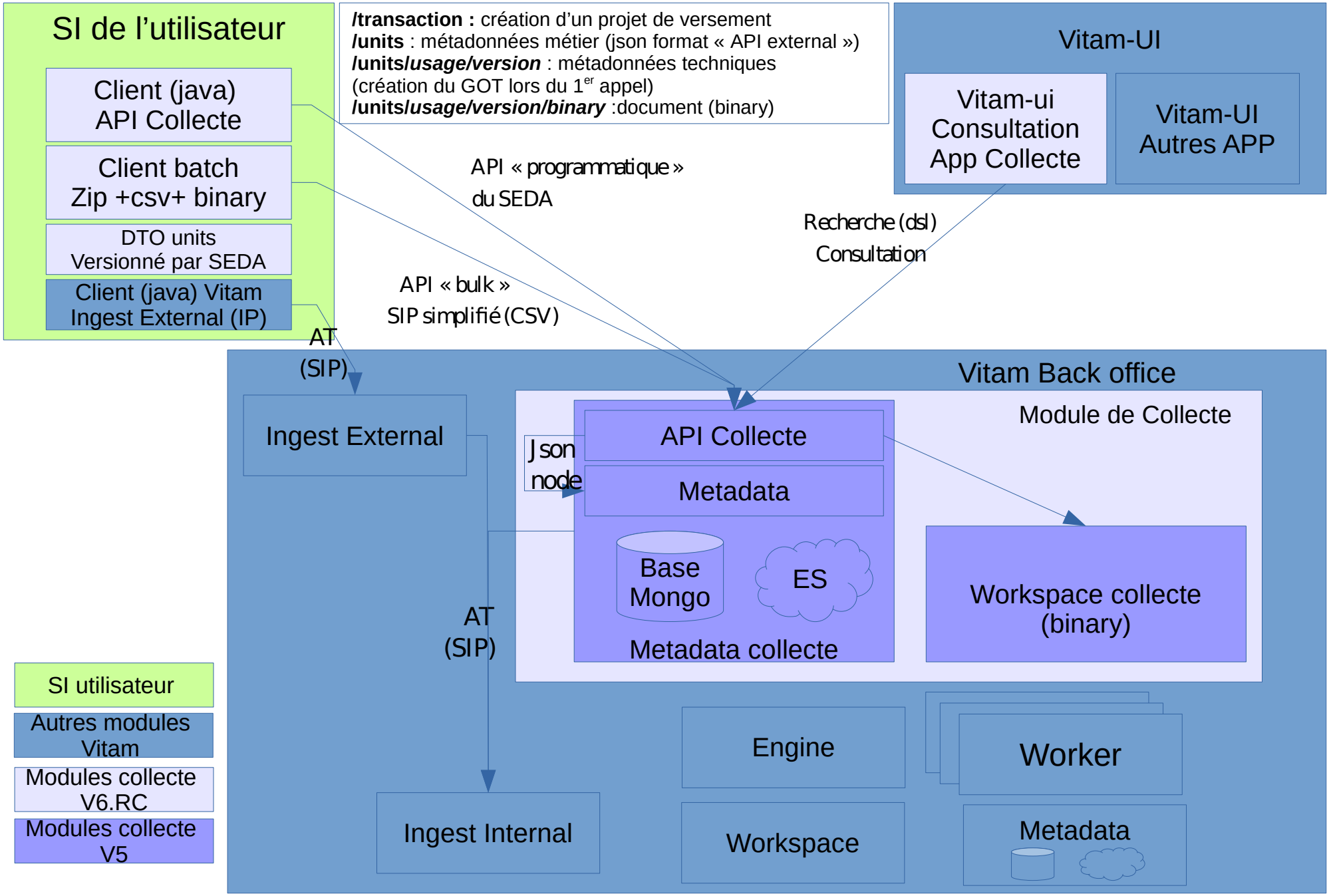
Développements et publications

- En **V5 (avril 2022)**, une première version bêta d'API de collecte en mode flux est disponible.

- En **V6.RC (octobre 2022)**, les écrans de traitement des versements bureautiques et un connecteur avec API en mode lot viendront compléter ce module.

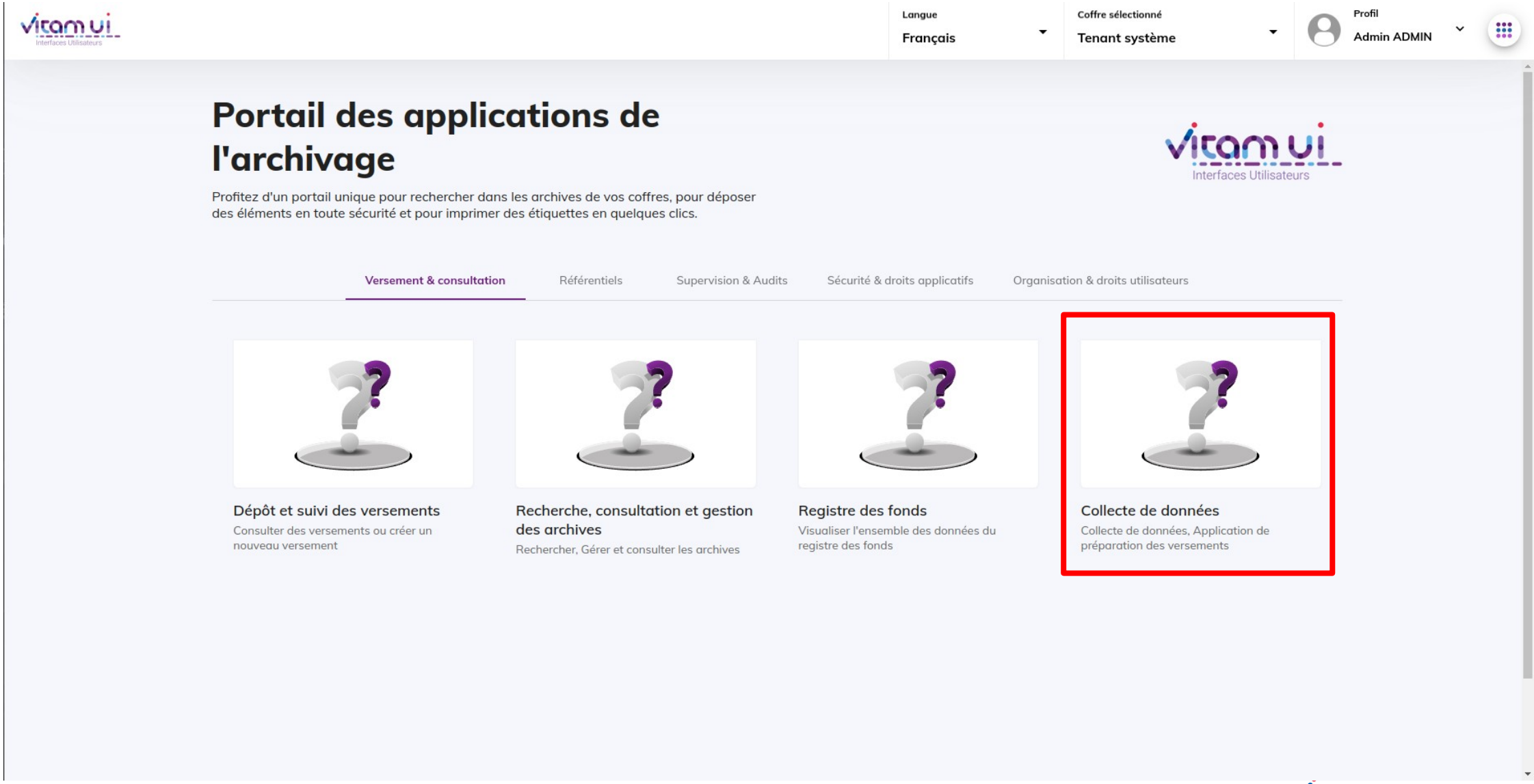
- Le module réutilise largement des composants internes afin de minimiser le coût de maintenance et permet d'envisager le développement de contrôles liés aux opérations d'entrée dans le module de collecte afin d'aider l'utilisateur à procéder à un versement conforme car déjà contrôlé.


VERSEMENT PAR FLUX
VERSEMENT MANUEL
Arborescences bureautiques
Autres types de documents
Système d'information de l'utilisateur
VitamUI
APP Collecte
SIP
Données + Fichiers
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


SI de l'utilisateur
Client (java)
API Collecte
Client batch
Zip +csv+ binary
DTO units
Versionné par SEDA
Client (java) Vitam
Ingest External (IP)
/transaction : création d'un projet de versement
/units : métadonnées métier (json format « API external »)
/units/usage/version : métadonnées techniques
(création du GOT lors du 1er appel)
/units/usage/version/binary :document (binary)
Vitam-UI
Vitam-ui
Consultation
App Collecte
Vitam-UI
Autres APP
API « programmatique »
du SEDA
Recherche (dsl)
Consultation
API « bulk »
SIP simplifié (CSV)
AT
(SIP)
Vitam Back office
Ingest External
Module de Collecte
Json
node
API Collecte
Metadata
Base
Mongo
ES
Metadata collecte
Workspace collecte
(binary)
AT
(SIP)
Engine
Worker
Ingest Internal
Workspace
Metadata
SI utilisateur
Autres modules
Vitam
Modules collecte
V6.RC
Modules collecte
V5

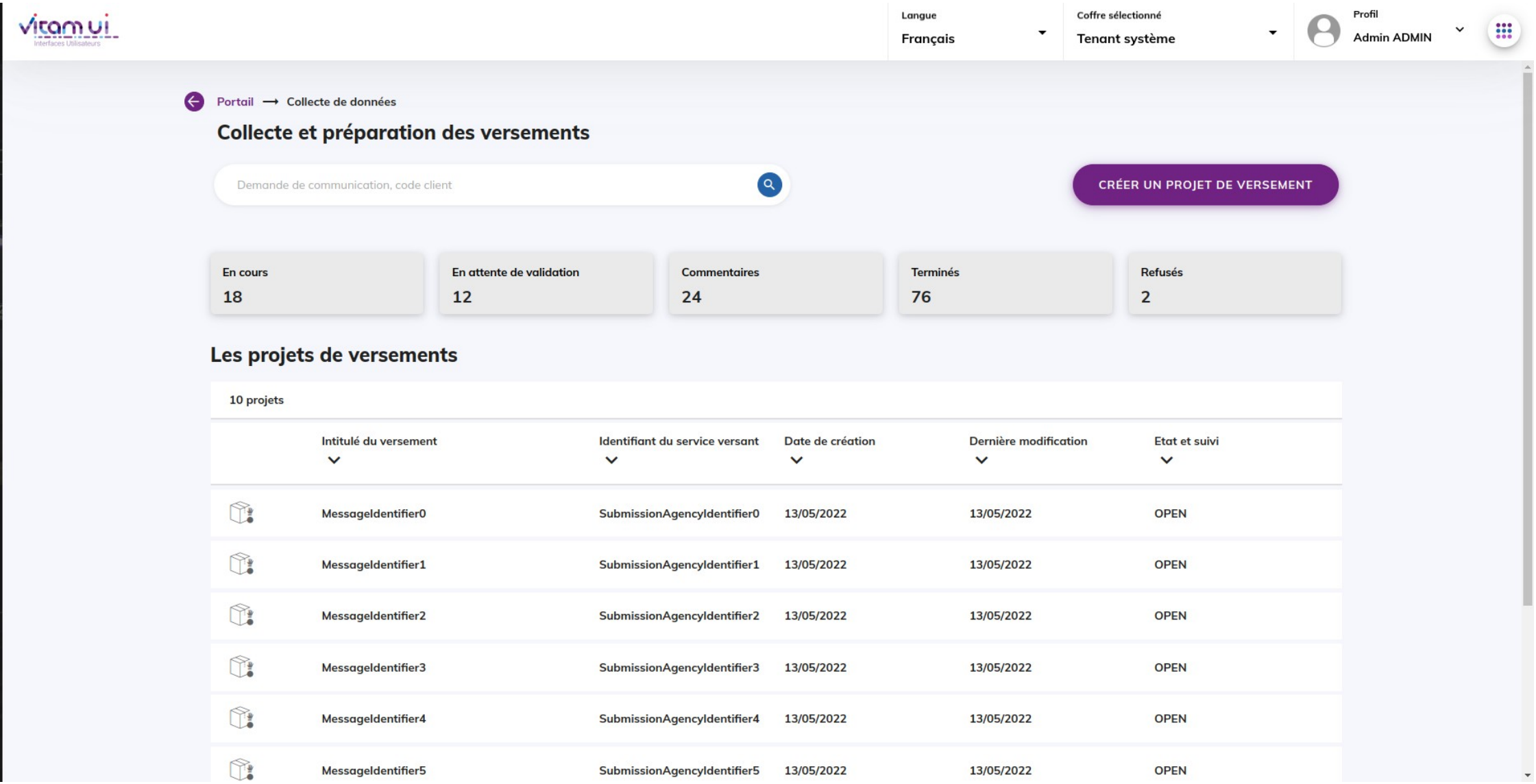
Vitam UI et la collecte d’archives bureautiques

Accès à l’APP Collecte des données



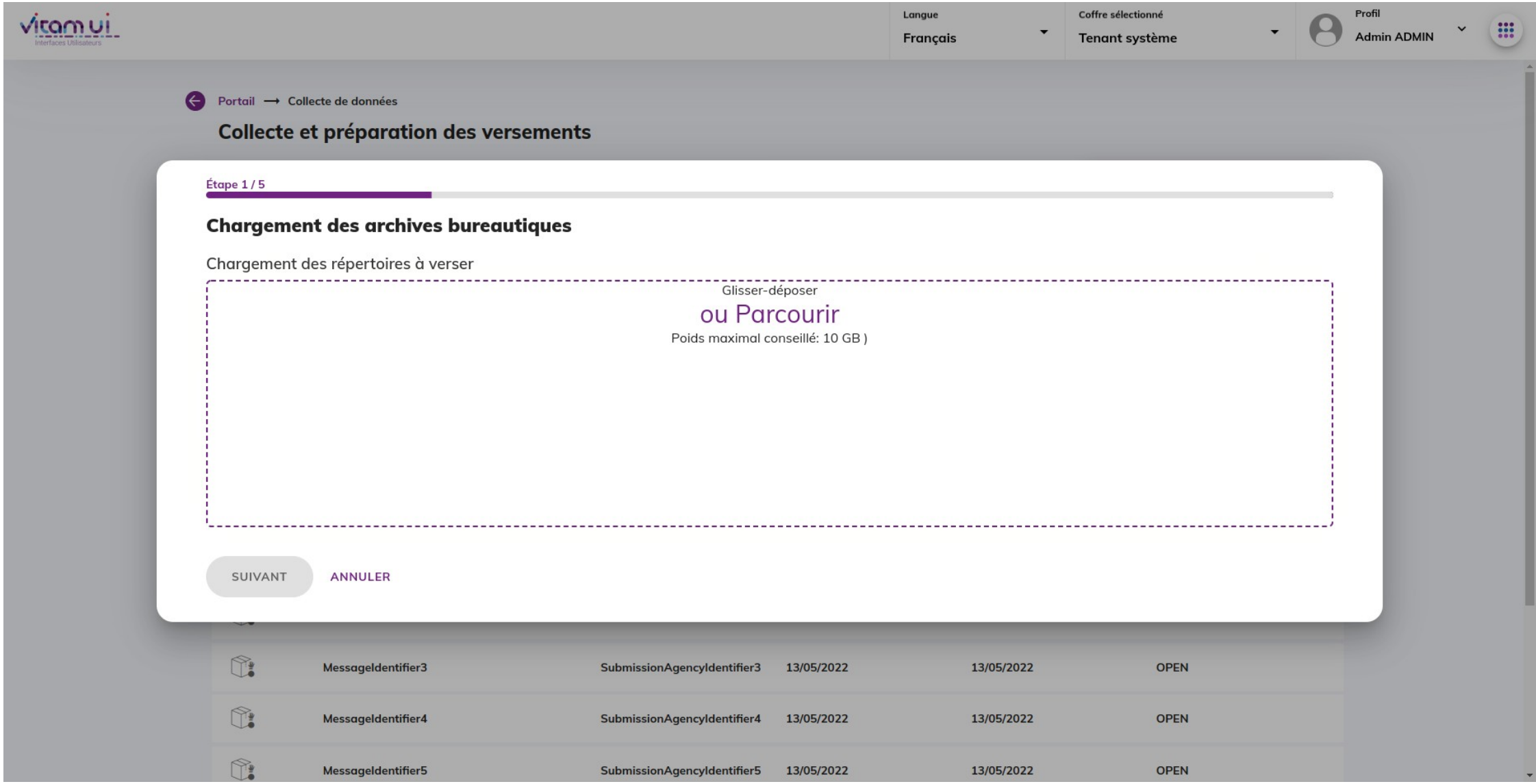
vitam
archivage numérique

Consultation de la liste des projets de versement



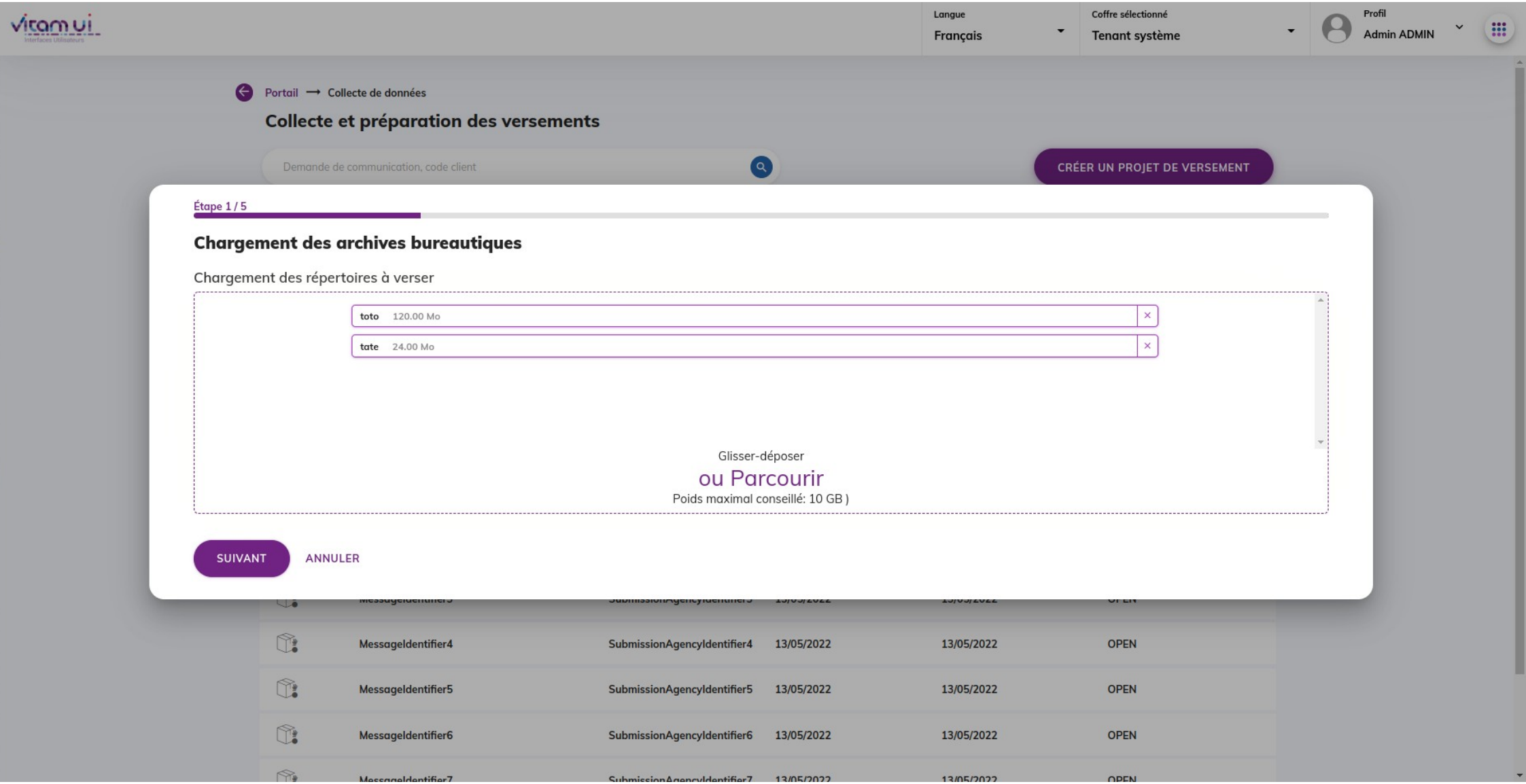
vitam
archivage numérique

Charger une arborescence bureautique (étape 1/5)



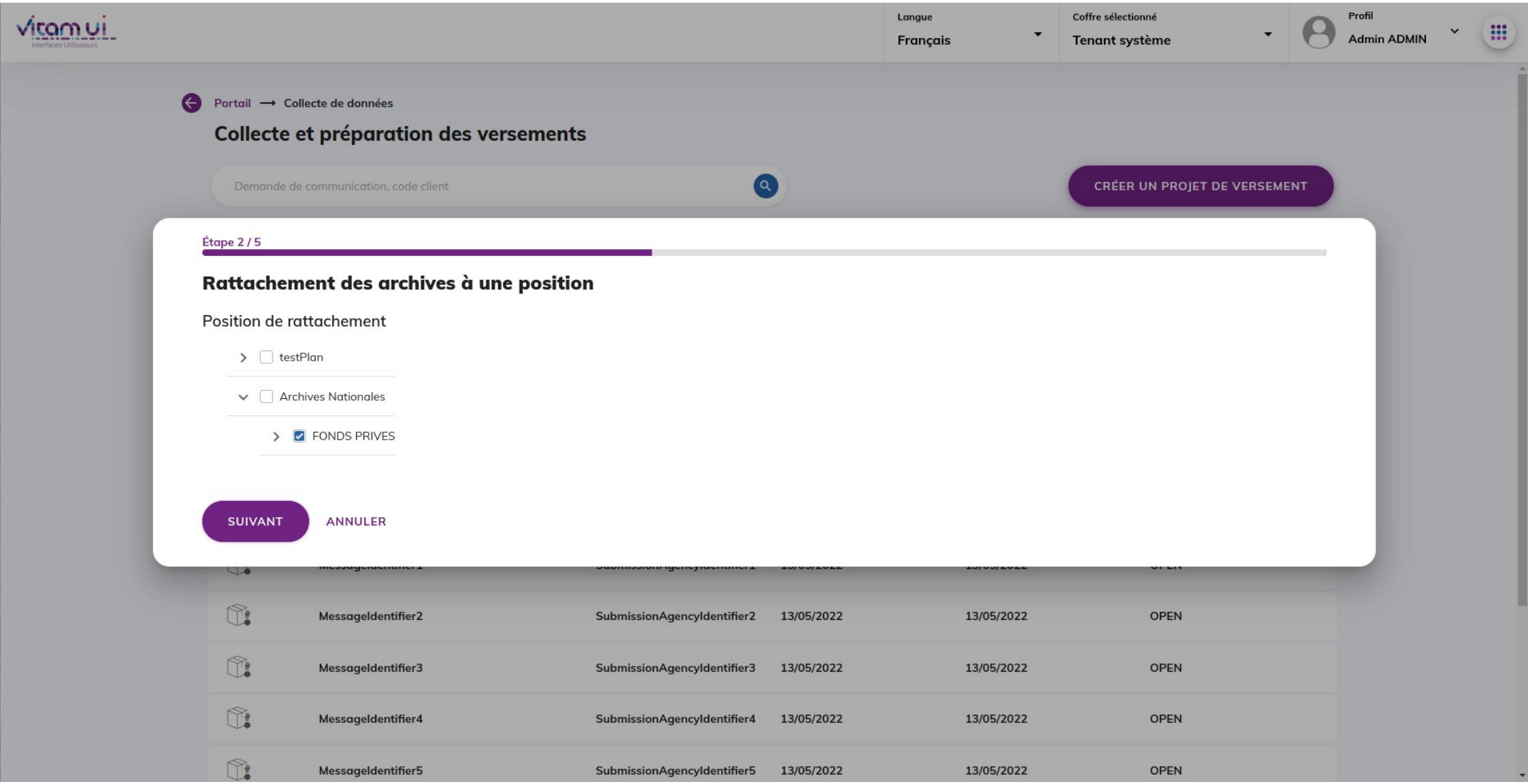
vitam
archivage numérique

Charger une arborescence bureautique (étape 1/5)



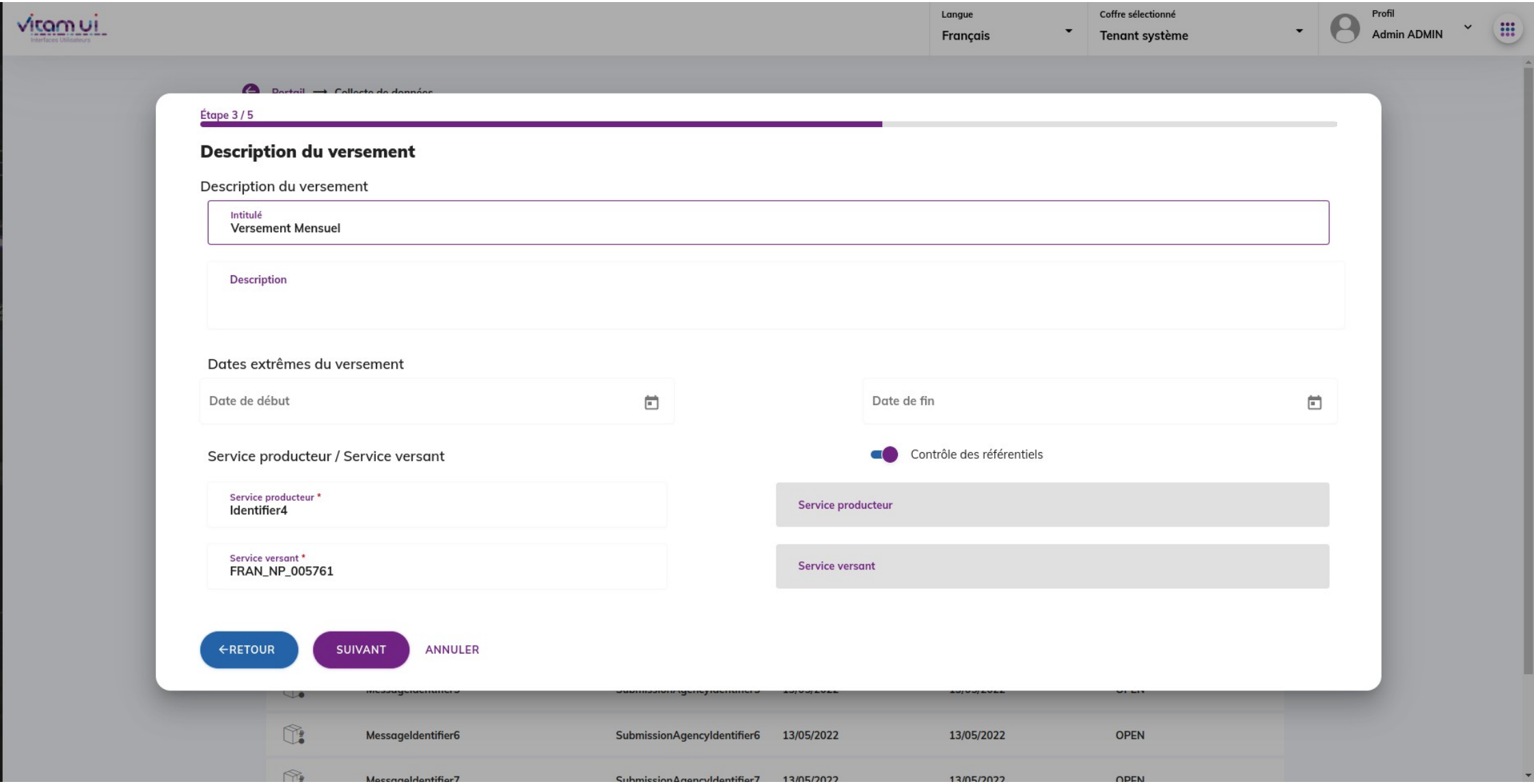
vitam
archivage numérique

Rattachement à une position (étape 2/5)

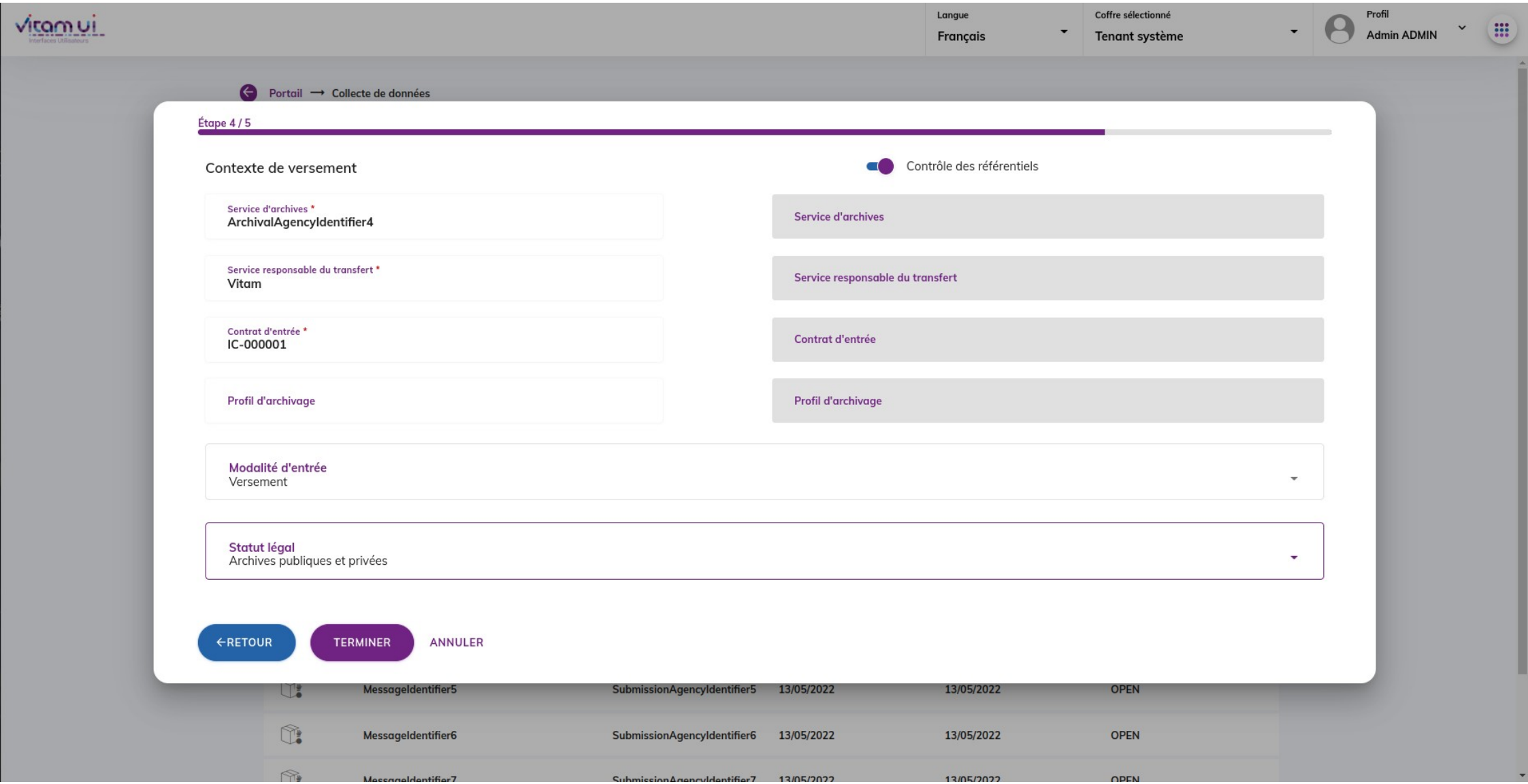


vitam
archivage numérique

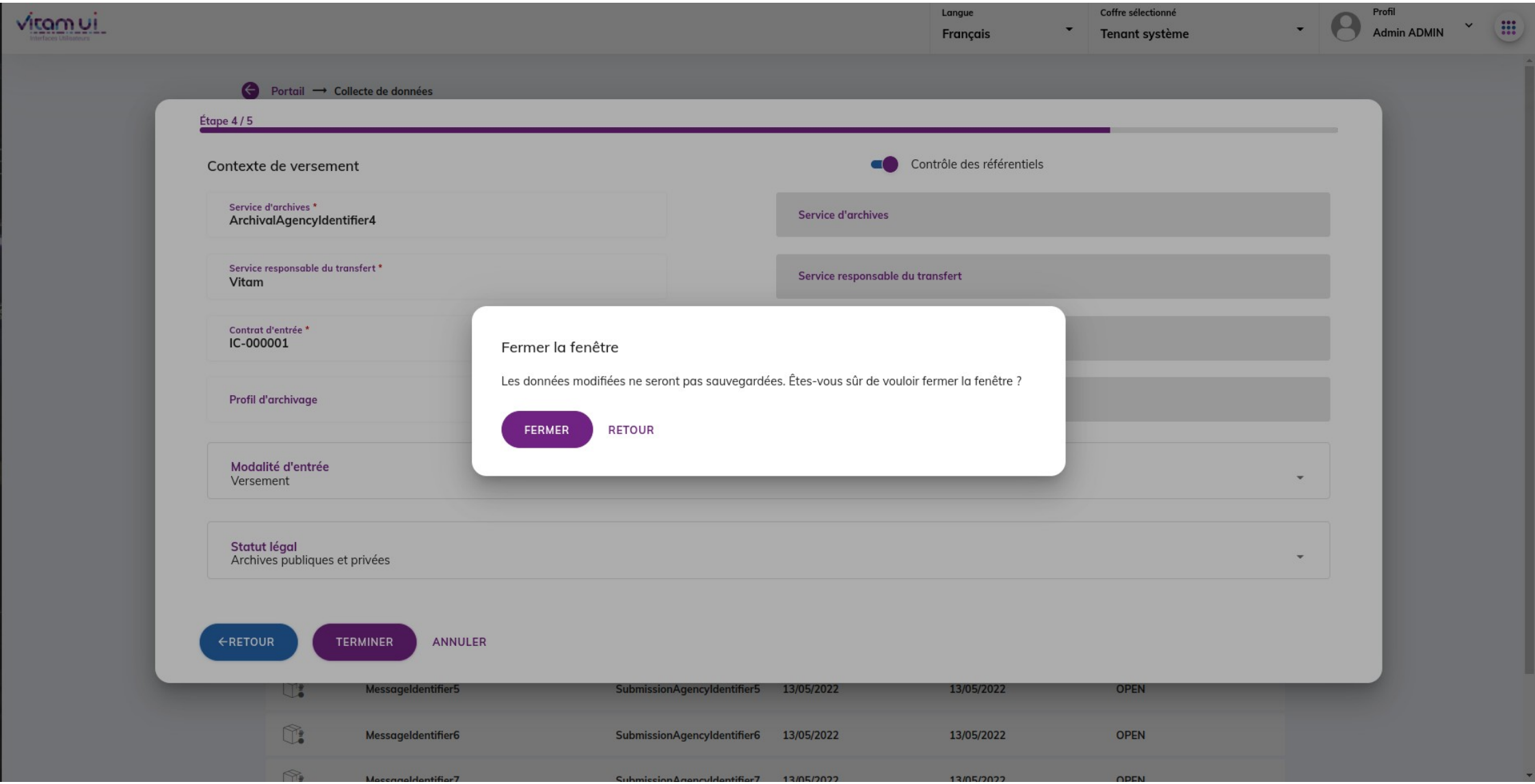
Description du versement (étape 3/5)



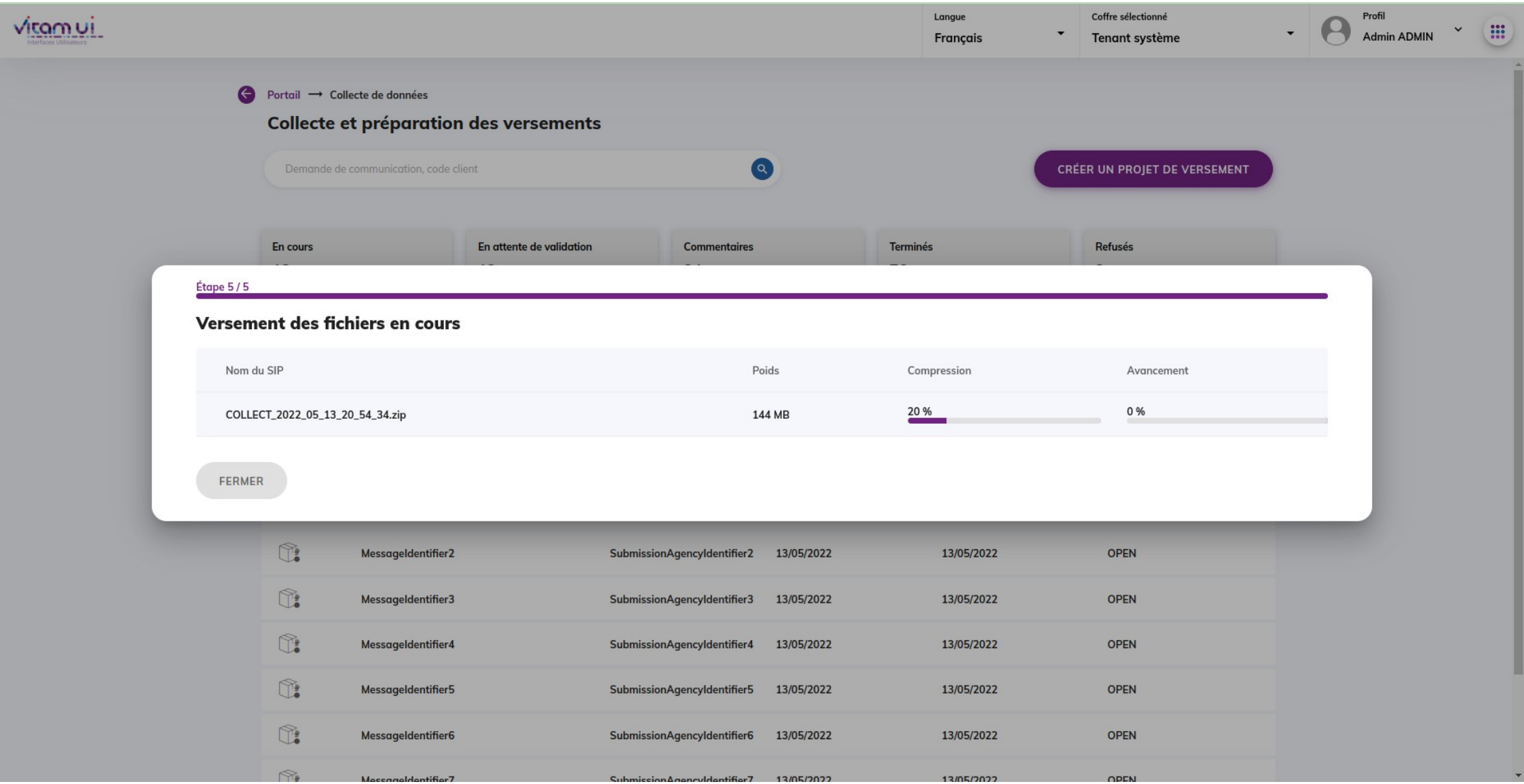
Contexte du versement (étape 4/5)



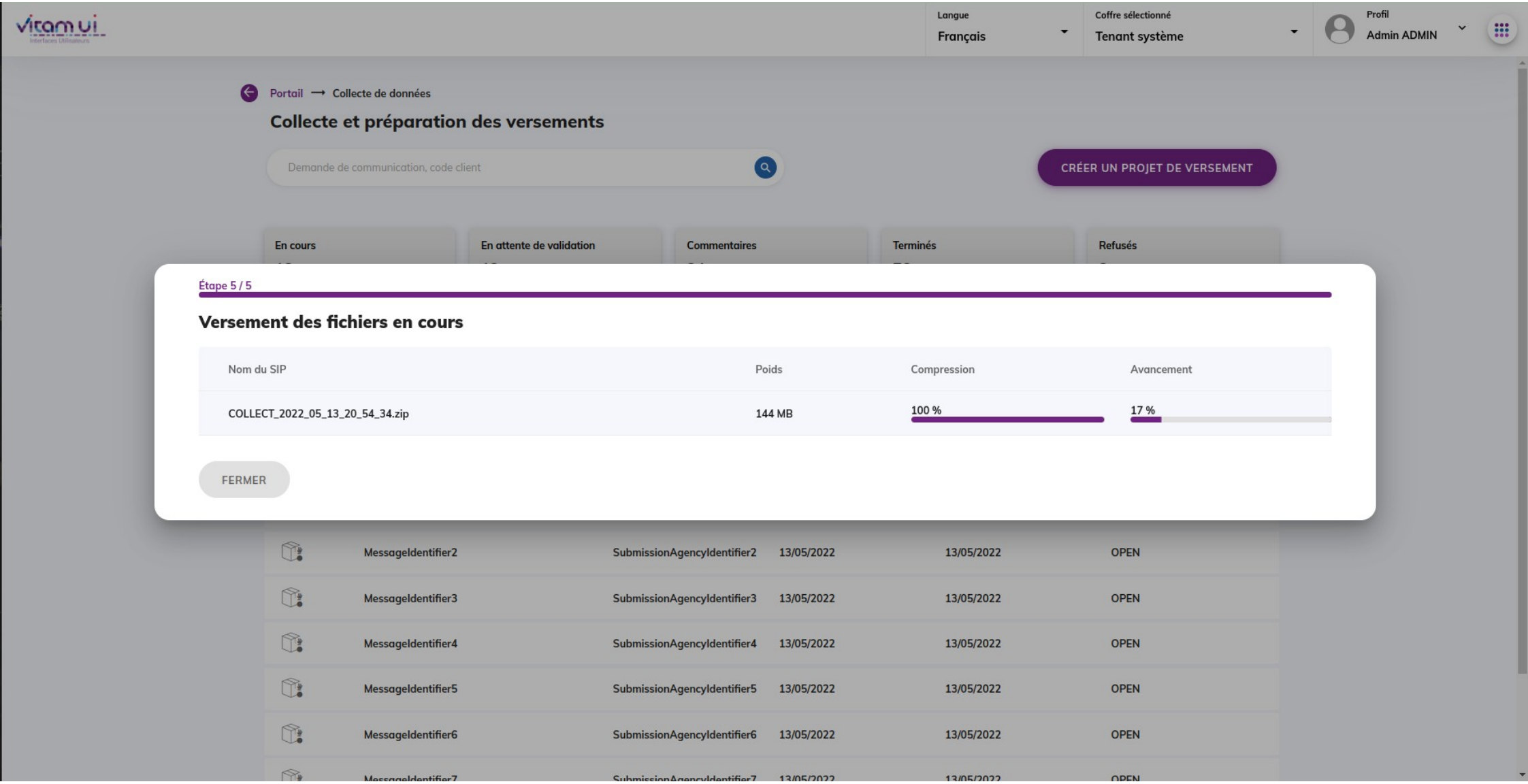
Confirmation de la création du projet (étape 4/5)



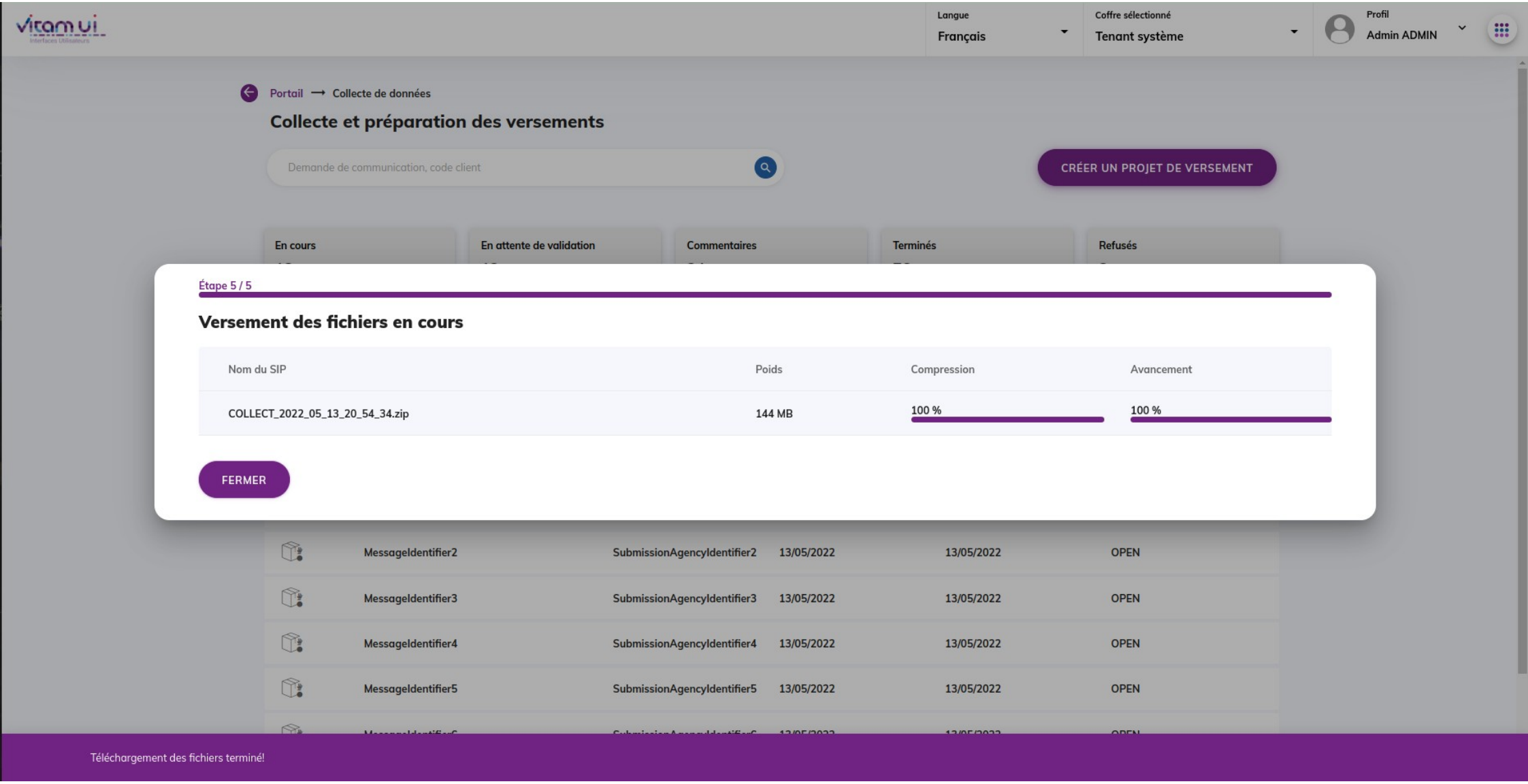
Enregistrement du projet de versement (étape 5/5)



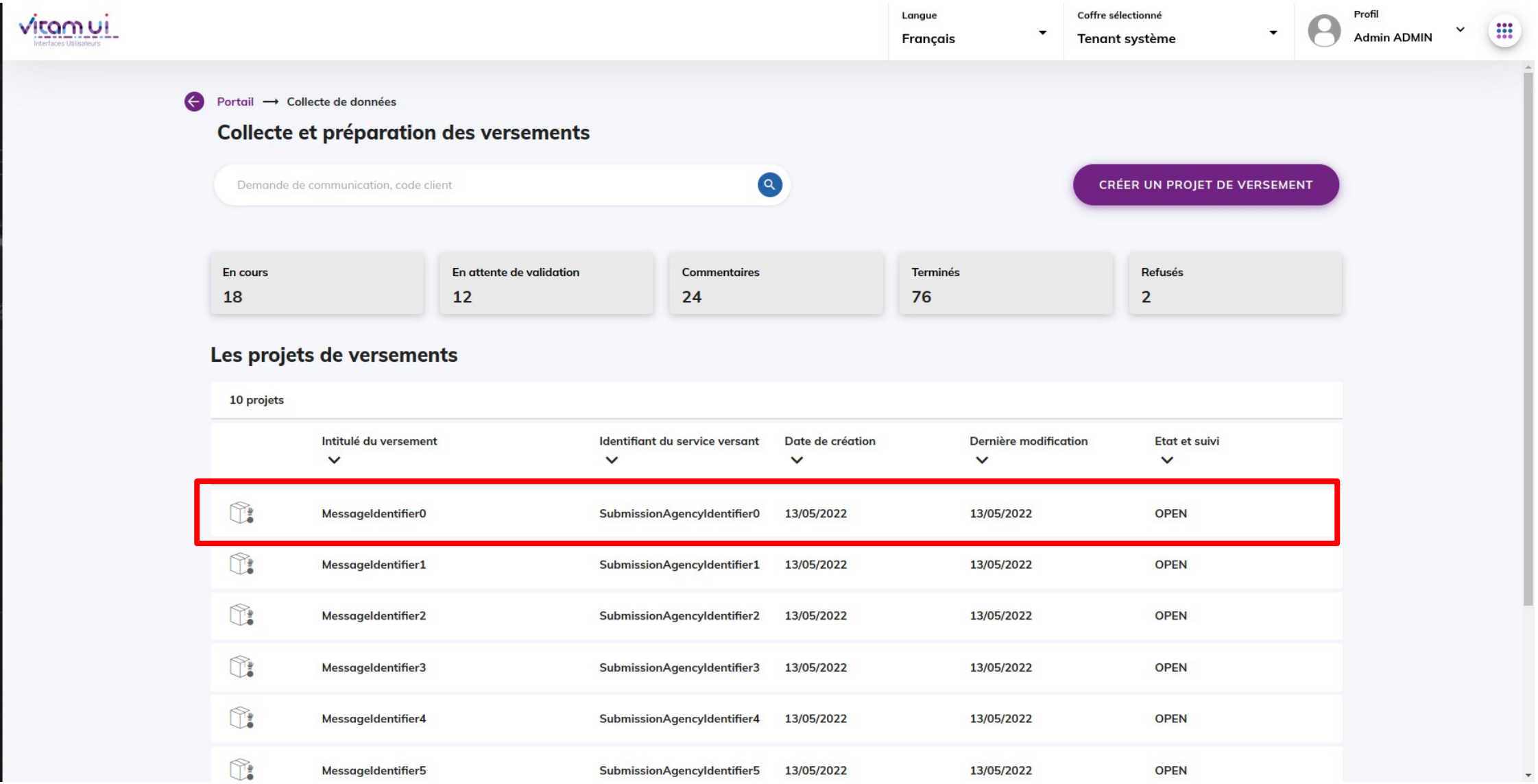
Enregistrement du projet de versement (étape 5/5)



Enregistrement du projet de versement (étape 5/5)

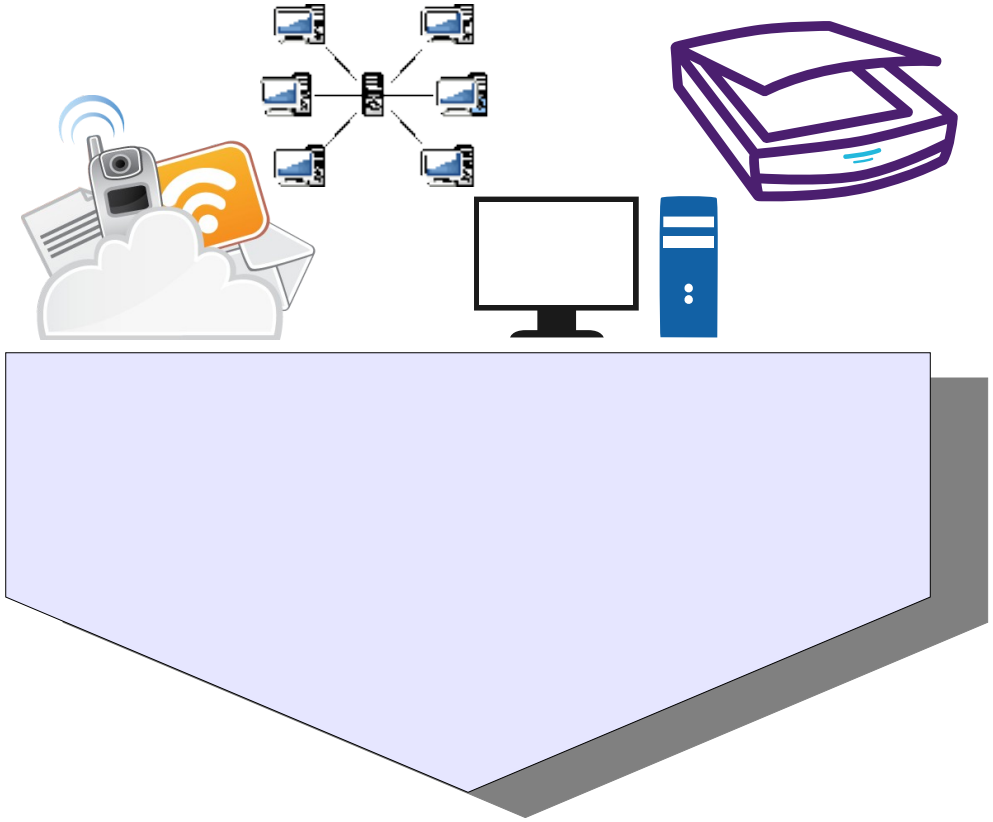


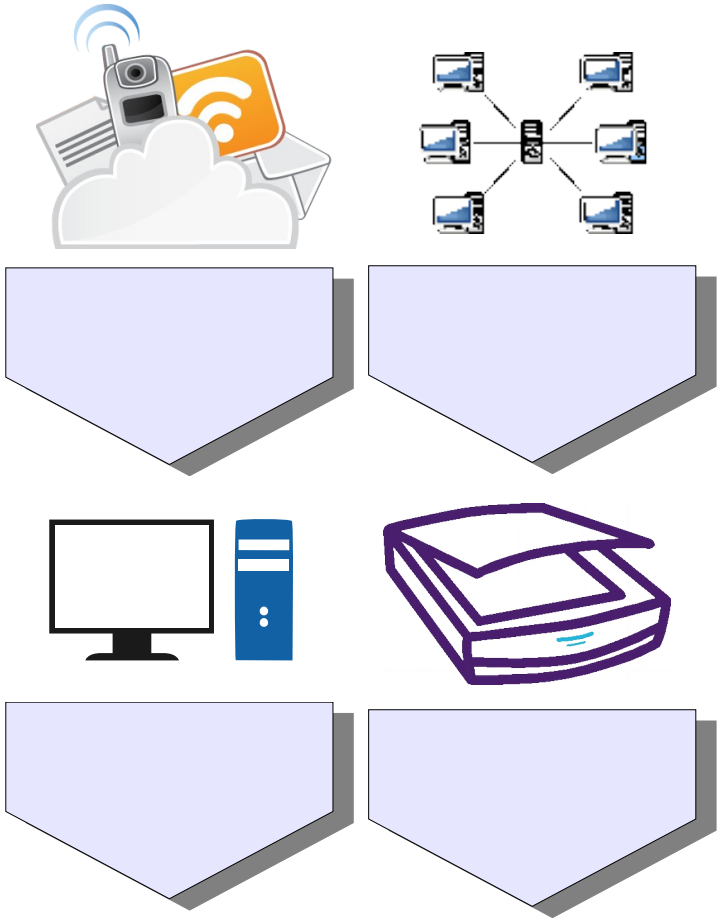
Enregistrement du projet dans la liste des projets de versement



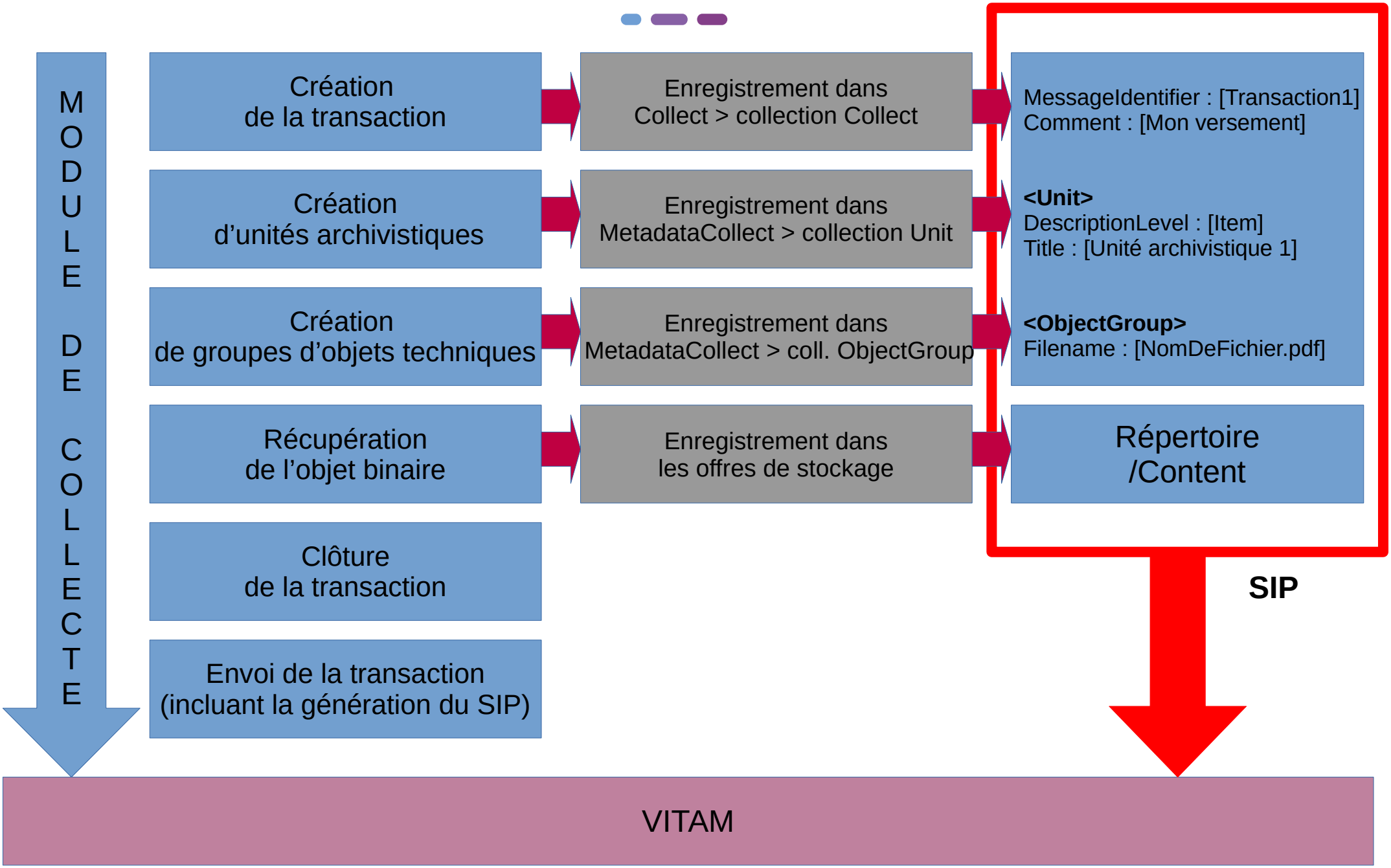
vitam
archivage numérique

Vitam et la collecte d’archives applicatives





Mécanisme mis en place



Démonstration d'un versement utilisant le module de collecte

En-tête


```
"ArchivalAgencyIdentifier": "Vitam",
"TransferingAgencyIdentifier": "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": ["aeeaaaaaaghefnffaaxa4ama6vfwaeqaaaaq"],
"#management": {
}
```


GOT


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


OT

ID909.pdf

vitam
archivage numérique

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

Des questions ?

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

PROGRAMME vitam

archivage numérique