



Ecole d'Ingénieurs Centrale Marseille

Sopra Groupe  
Aix-en-Provence, France

## Travail de Fin d'Etudes

Du Lundi 31 Mars 2014 au Vendredi 26 Septembre 2014

# Projet de Refonte du SIRH de la Ville de Marseille AZUR SIRH

## Rapport de Stage

Bach Phan Luong

3<sup>ème</sup> Année, promotion 2014  
Parcours OMIS – Informatique

Tuteur école :  
François Brucker

Tuteur Entreprise :  
Bernard Spatz

# Sommaire

---

Remerciements	4
Résumé	5
In Brief	5
Mots – clés	6
Glossaire / Abréviation	7
1. Introduction	8
2. L'entreprise	9
<b>2.1. Domaine d'activités et chiffres</b>	<b>9</b>
<b>2.2. Organigramme et gestion des ressources</b>	<b>12</b>
2.2.1. Organigramme	12
2.2.2. Organisation des ressources	12
2.2.3. Projets en régie et au forfait	12
3. Le Projet	13
<b>3.1. Le contexte du projet</b>	<b>13</b>
<b>3.2. Les phases du projet</b>	<b>14</b>
<b>3.3. Les modules du projet</b>	<b>15</b>
4. Les missions confiées	17
<b>4.1. Le module Editique</b>	<b>17</b>
<b>4.2. Le module Reprise de données</b>	<b>18</b>
5. Le module Editique	19
<b>5.1. La problématique</b>	<b>19</b>
<b>5.2. Proposition théorique de résolution</b>	<b>21</b>
5.2.1. Adressage des Éditions	22
5.2.2. Aiguillage fonctionnel	23
5.2.3. Extraction et constitution d'un XML de données	23
5.2.4. Editique Starjet APPIC	23
<b>5.3. Mise en place de la solution</b>	<b>24</b>
5.3.1. Constitution d'une édition	24
5.3.2. Prise en charge côté APPIC	32
6. Le module Reprise de données	37
<b>6.1. La problématique</b>	<b>37</b>

<b>6.2. Proposition théorique de résolution</b>	<b>38</b>
6.2.1. Schéma global : BDD source – ETL Talend – BDD cible	38
6.2.2. Reprendre chaque entité par un job Talend	40
6.2.3. Principe de déclenchement des jobs	40
6.2.4. Attendus supplémentaires : Journaux de rejets, de statistiques et d'erreurs d'exécution	40
<b>6.3. Mise en place de la solution</b>	<b>41</b>
6.3.1. L'organisation d'un job Talend	41
6.3.2. La définition du schéma de données par des méta-données	42
6.3.3. Les traitements spécifiques	44
6.3.4. L'intégration d'exécutables complémentaires	47
6.3.5. La variabilisation des chemins et noms de fichiers par des variables de contexte	47
6.3.6. Intégration dans Pléiades	48
<b>6.4. Qualification du travail</b>	<b>49</b>
6.4.1. L'outil : Quality Center HP	49
6.4.2. Le principe de Qualification	50
<b>7. Conclusion</b>	<b>51</b>
<b>8. Annexes</b>	<b>52</b>
<b>8.1. La Fonction Publique</b>	<b>52</b>
<b>8.2. Les formations reçues dans le cadre du projet</b>	<b>53</b>
8.2.1. La solution éditique APPIC	53
8.2.2. Le module Pléiades de gestion de la paie et des temps d'activité PTA	53
8.2.3. L'ETL Talend	53
8.2.4. L'outil de contrôle de tests Quality Center HP	53
8.2.5. La solution base de données RH de Sopra : Pléiades	53
<b>8.3. Autres connaissances mobilisées</b>	<b>54</b>
8.3.1. Les macros VBA	54
8.3.2. Le langage Java	54
8.3.3. La conception MCD Modèle Conceptuel de Données	54
8.3.4. Les connaissances en SQL	54
<b>8.4. Cycle de vie en V incluant les tests</b>	<b>55</b>
<b>8.5. Principales Références Clients de la Division Sud Est</b>	<b>56</b>
<b>8.6. Bibliographie</b>	<b>57</b>

## Remerciements

---

La réalisation de ce stage de fin d'étude est une étape essentielle à la construction de mon parcours professionnalisant dans le cadre de mes études à Centrale Marseille. Elle a été rendue possible grâce à la confiance que m'avait accordée Frédéric Liccia, le Directeur de l'agence 121 Sopra Aix, ainsi que Olivier Vandembroucke, Senior Engineer et responsable technique à l'agence. Je leur dois une sincère reconnaissance.

Durant mon stage de fin d'étude, j'ai été accompagné par mon tuteur Bernard Spatz, Directeur de projets. Malgré toutes les charges de travail que lui imposent son poste et ses responsabilités, il a été d'une grande patience et d'une grande disponibilité aux problèmes que je rencontrais. Grâce à lui, j'ai été intégré à l'équipe avec une responsabilité pleine et entière au même titre qu'un jeune collaborateur. Je souhaite le remercier vivement de son implication et de son accompagnement durant mon stage.

Par ailleurs, l'intérêt d'un stage ou plus globalement d'une vie professionnelle ne réside pas seulement dans l'intérêt du travail ou des missions effectuées. L'ambiance de travail et les relations humaines sont aussi très importantes. Ainsi je souhaite saluer chaleureusement tous les membres de l'équipe de mon projet, notamment ceux de mon bureau, Sami Roukhi, Kamlane Tsuen, Solange Gastinel, Guillaume Kuhmel et tout particulièrement Jérémy Soria pour l'intégration dont j'ai pu en bénéficier ainsi que l'esprit d'entraide et le travail d'équipe que nous avons pu forger.

Ce stage de fin d'étude étant l'aboutissement du cursus d'école d'ingénieurs de Centrale Marseille, il est donc le fruit de ces trois années de formation à l'école. Je souhaite donc remercier vivement mes professeurs et les responsables pédagogiques et administratifs pour leur dévouement et le travail formidable qu'ils réalisent quotidiennement dans l'intérêt de l'école et des élèves.

## Résumé

---

Ce stage de fin d'étude m'a donné l'opportunité de découvrir l'organisation fonctionnelle, l'environnement de travail ainsi que les activités de Sopra à travers son agence 121 Aix. Le groupe Sopra est une ESN d'envergure européenne spécialisée en Conseil, en Services Technologiques et en Edition de logiciels.

J'ai ainsi pu participer aux différentes phases des modules Editique et Reprise de données du projet de Refonte du SIRH de la Ville de Marseille. Concernant le module Editique, il s'agit d'automatiser le processus d'impression de documents officiels sur la base de règles de gestion du client. La Reprise de données consiste en le transfert des données du SIRH d'origine vers le nouveau SIRH, tout en respectant également des règles de gestion.

Aussi j'ai pu développer des compétences techniques, d'analyse-conception, de développement et de rédaction, et également des méthodologies de travail et d'organisation ainsi que des qualités relationnelles et communicationnelles.

La participation à ce projet me permet donc à la fois de développer des compétences fondamentales en Informatique et à la fois d'avoir une première approche sur un projet de refonte d'un SI de grande ampleur.

## In Brief

---

This internship has given me the opportunity to discover the functional organization, the working environment and also the various activities of Sopra by my daily work at the agency Aix 121. Sopra Group is a European-wide IT services company specialized in IT Consulting and Services and in Software Edition.

Indeed, I'm participating in the various phases of modules Publishing and Data Transfer of the project of Reorganisation of the HRIS of the City of Marseille. Regarding the Publishing module, it aims to automate the process of printing official documents based on Human Resources rules of the client. Data transfer consists of data transferring from the original HRIS to the new HRIS, while also respecting HR rules.

Then, I have the opportunity to develop technical skills, analysis and design, development and specification writing, and also methodologies of work and organization as well as relational and communicational qualities.

Participation in this project allows me to both develop fundamental basic skills in computer science and to have an idea on a project of reorganization of a large-scale IT infrastructure.

## Mots – clés

---

Informatique de Gestion

Refonte du SIRH de la Ville de Marseille

Système d'Information en Ressources Humaines

Pléiades, solution HR Sopra

Editique

Star Jet APPIC

Reprise de données

ETL Talend

Qualification

Quality Center

VBA Excel

Stage de fin d'étude, 3A Centrale Marseille

Sopra Groupe

## Glossaire / Abréviation

---

*ESN* : Entreprise en Service Numérique

*CA* : Chiffres d'Affaires

*BDD* : Base de Données

*SGBD* : Système de Gestion de Base de Données

*Le client* : La Ville de Marseille

*SI* : Système d'Information

*SIRH* : Système d'Information en Ressources Humaines

*RH* : Ressources Humaines

*ETL* : Extract Transform Load

*SJD* : StarJet Design

*FP* : Fonction Publique

*SVN* : Outil de versionning

*TMA* : Tierce Maintenance Applicative

*CSM* : Centre de Service Mutualisé. Le CSM peut intervenir sur tous les projets de l'agence à laquelle il est affecté.

*CRM* : Customer Relationship Management. Il s'agit d'un pôle de compétence de l'agence Sopra 121.

*Agence 121* : Il s'agit de l'agence Sopra Provence Business Technologies sur le site d'Aix-les-Milles.

*Projet en régie* : Projet dont la facturation est basée sur le temps réel mis par le prestataire de service.

*Projet au forfait* : Projet dont la facturation est basée sur une évaluation ferme et définitive lors de la phase de négociation initiale. Des avenants peuvent cependant être négociés en cas d'imprévu.

*Projet clé en main* : Projet dont le livrable est immédiatement prêt à être utilisé.

## 1. Introduction

---

Participer à un projet d'envergure me permettant d'avoir une vue d'ensemble et de prendre part aux différentes phases d'un projet informatique était l'un de mes critères pour choisir mon stage. Mon stage de fin d'étude m'en offre l'opportunité. Il s'agit en effet d'un stage de fin d'étude de six mois du cursus de l'école d'ingénieurs Centrale Marseille, avec une spécialisation en informatique, dont l'objectif est de mettre le stagiaire en situation réelle d'un ingénieur et de pouvoir mobiliser ses connaissances acquises à l'école.

Mon stage s'inscrit ainsi dans un projet de grande ampleur portant sur trois ans consistant en la refonte du SIRH d'une grande communauté urbaine, la Ville de Marseille.

Ce projet est mené par Sopra, une ESN, mon entreprise d'accueil.

Ainsi ce rapport de stage décrira tout d'abord l'entreprise et son activité.

Il présentera ensuite avec une vue d'ensemble le projet auquel je participe.

Je présenterai ensuite les missions confiées pour terminer avec l'analyse technique des solutions mises en œuvre.



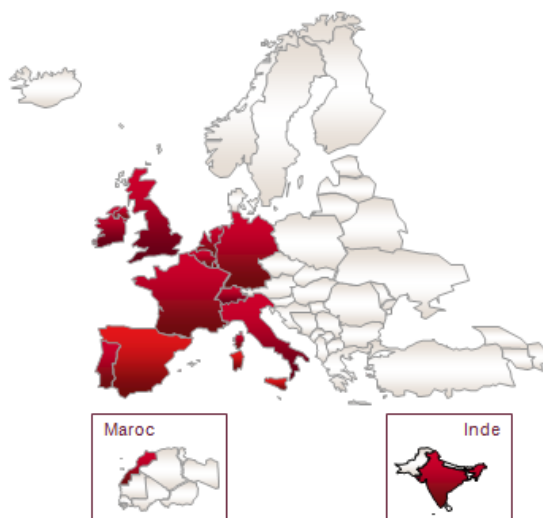
## 2. L'entreprise

### 2.1. Domaine d'activités et chiffres

#### *Le groupe Sopra*

Le groupe Sopra est une ESN en Conseil (SI et Management), en Services Technologiques (Intégration de SI et Outsourcing applicatif) et en Edition de logiciels (Banque, Immobilier, RH). C'est un acteur majeur indépendant qui emploie plus de 16000 collaborateurs dans le monde pour un CA d'environ 1.349 milliards d'euros en 2013.

- Groupe européen indépendant
- Création en 1968
- Top 10 des Sociétés de Conseil et de Services en Technologies de l'Information en Europe
- Les 3 métiers de Sopra:
  - Conseil (7%)
  - Services Technologiques (72%)
  - Edition de logiciels (21%)



Sopra est un groupe en bonne croissance, dont voici les résultats sur les trois dernières années (2011 à 2013).



De plus, Sopra effectue actuellement un rapprochement avec Steria à travers une OPE amicale sur la base d'une action Sopra pour quatre actions Steria. (source : [http://www.sopra.com/docs/default-source/Documents-financiers/30062014\\_fr\\_soprasteriaouverture\\_ope\\_final.pdf?sfvrsn=0](http://www.sopra.com/docs/default-source/Documents-financiers/30062014_fr_soprasteriaouverture_ope_final.pdf?sfvrsn=0) )

(La source interne confidentielle n'est pas présentée ici).

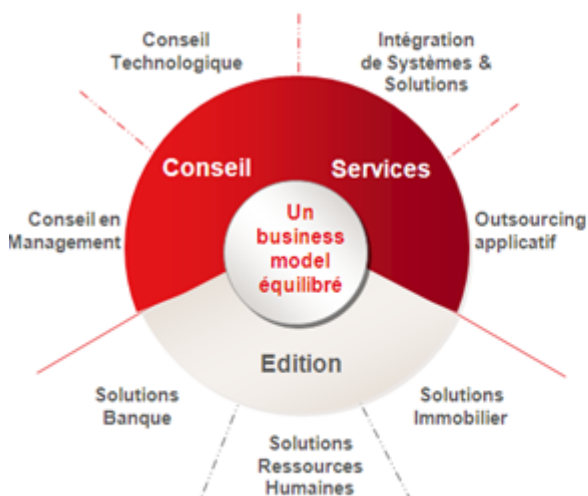
Le résultat de la fusion donnera naissance à un groupe d'envergure mondial pesant plus de 3,1 Mds € DE CA, pour plus de 35 000 collaborateurs dans 24 pays dans le monde.

(source : <http://www.steria.com/fr/medias/communique/article/projet-de-rapprochement-amical-de-sopra-et-de-steria-pour-creer-un-leader/>)

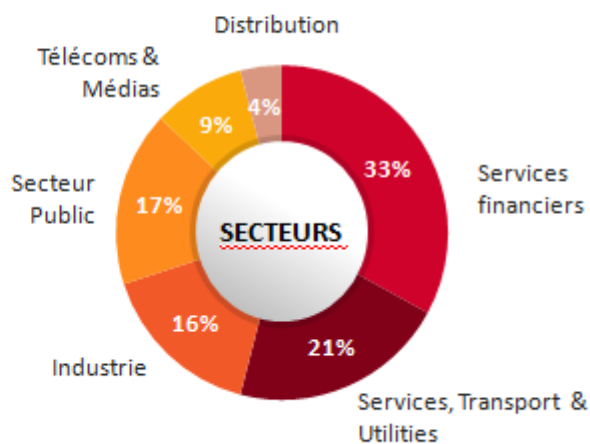
Cependant la suite de la présentation de Sopra sera focalisée sur l'activité de Sopra en 2013.

La répartition de l'activité de Sopra est la suivante.

- **Un continuum de services** pour accompagner durablement les entreprises dans leurs projets de transformation



- **Sopra leader** auprès des grands donneurs d'ordre en France et en Europe



### La division Sud-Est



La division Sud-Est de Sopra est composée de quatre sites géographiques : Montpellier, Aix, Toulon et Sophia-Antipolis.

Le groupe y réalise un CA de 65 millions en 2013 et y emploie environ 750 collaborateurs.

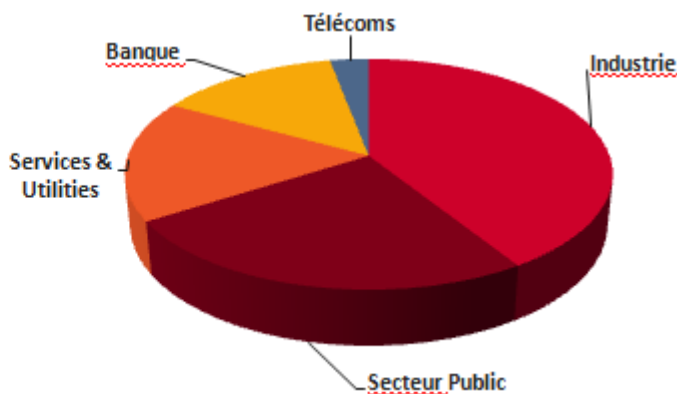
### L'agence Provence Business Technologie

Le site géographique d'Aix-les-Milles est constitué de quatre agences, Aeroline, Conseil, STIE et Provence Business Technologie, auquel je suis rattaché.

La répartition de l'activité de l'agence est schématisée ci-dessous.



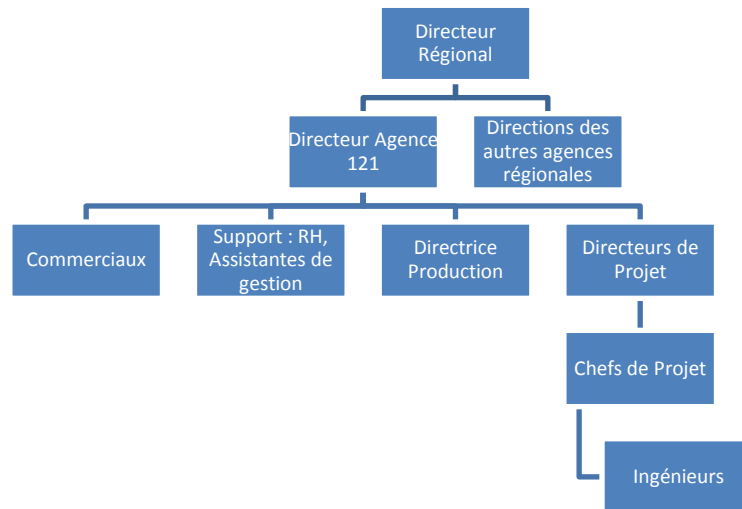
**235 collaborateurs**  
**20 millions d'euros de CA**  
**50 comptes actifs**  
**70 % de l'activité au forfait**



Vous retrouverez en annexes les principales références de la division Sud-Est.

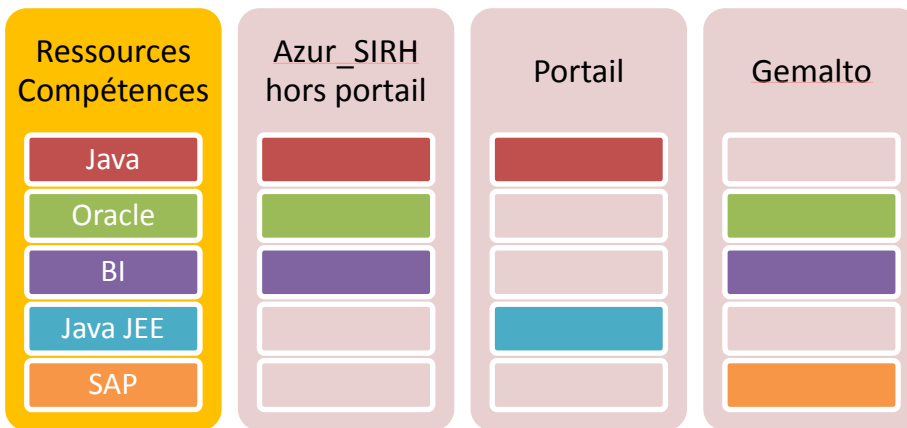
## 2.2. Organigramme et gestion des ressources

### 2.2.1. Organigramme



### 2.2.2. Organisation des ressources

Pour simplifier, l'organisation des ressources de l'agence 121 suit une structure matricielle. Une illustration est donnée sur trois projets gérés par l'agence.



A cela s'ajoute des centres de service comme le CMS Qualité-Testing ainsi que des pôles de compétence comme le pôle CRM.

Enfin, l'affectation des ressources à chaque projet est décidée par la directrice de la Production.

### 2.2.3. Projets en régie et au forfait

Sopra réalise les projets avec le client en facturant ses prestations soit en régie soit au forfait.

Les projets au forfait peuvent être généralement des projets clé en main. Ils sont généralement réalisés sur les sites de Sopra. Le projet Refonte du SIRH de la Ville de Marseille est un projet au forfait.

Les projets en régie peuvent être généralement des évolutions applicatives ou des TMA. Ils sont généralement réalisés sur les sites du client.

## 3. Le Projet

---

### 3.1. Le contexte du projet

Les missions auxquelles je participe durant mon stage s'inscrivent dans le cadre du projet de refonte du SIRH de la Ville de Marseille et des prestations associées de mise en œuvre.

Le client, la Ville de Marseille, est donc un grand compte du secteur public ; une synthèse de la fonction publique est disponible en annexes. Remarquons qu'il est déjà client de Sopra sur d'autres projets comme les projets de portail web et d'application de gestion des élections ou du recrutement.

La gestion des ressources humaines du client est découpée en domaines fonctionnels, notamment la gestion de la carrière, la gestion prévisionnelle des emplois et des compétences (GPEC), la gestion de la paie, la gestion du temps d'activité (GTA), la gestion des absences et des congés, etc.

Le client possède actuellement un SIRH composé d'une centaine d'applications spécifiques et indépendantes, c'est-à-dire réalisées sur-mesure par différents prestataires. Leur utilisation et leur maintenance simultanée ne sont plus satisfaisantes.

Aussi la Ville de Marseille souhaite disposer d'un outil de pilotage des Ressources Humaines unique, souple, transversal, et optimisé autour des fonctionnalités globales en mesure d'intégrer les évolutions réglementaires, et adapté aux spécificités d'une collectivité qui emploie environ 12 000 agents exerçant plus de 200 métiers. L'ambition de la Ville est aussi, à terme, de faire évoluer les modes d'organisation de la fonction RH et des processus associés.

Basée sur sa solution RH Pléiades, acquis en 1996, l'enjeu pour Sopra consiste dans la conception, le déploiement et la maintenance d'un outil de pilotage stratégique et opérationnel couvrant l'ensemble du périmètre RH ainsi que la réalisation de trois portails dotés d'une interface conviviale, destinés respectivement aux agents, aux experts et aux managers de la Ville, soit un total de 800 utilisateurs.

Les technologies les plus récentes (coffre-fort électronique, signature numérique, GED) y sont intégrées et la sécurité des données garantie.

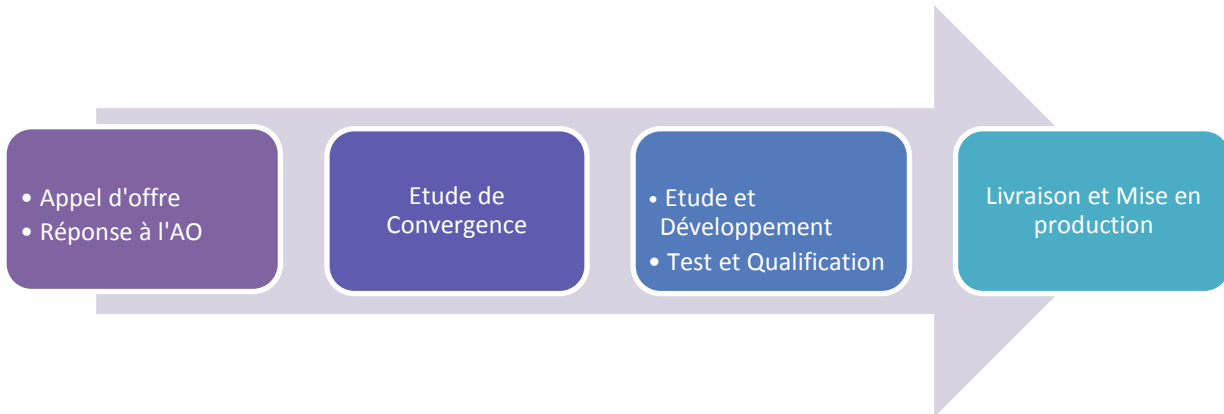
Ce projet de grande transformation se décline en 2 parties :

- La mise en œuvre progressive de l'ensemble du périmètre global RH, pour lequel Sopra, réalisera aussi l'intégration fonctionnelle et technique au sein du Système d'Information de la Ville de Marseille.
- La maintenance applicative de la solution, intégrant la veille légale et ses évolutions. Sopra et la Ville de Marseille s'inscrivent dans une démarche commune d'anticipation et d'adaptation aux diverses évolutions (techniques, technologiques, organisationnelles, ...).

La suite de ce rapport se limitera au périmètre de la réalisation de la refonte du SIRH de la Ville de Marseille.

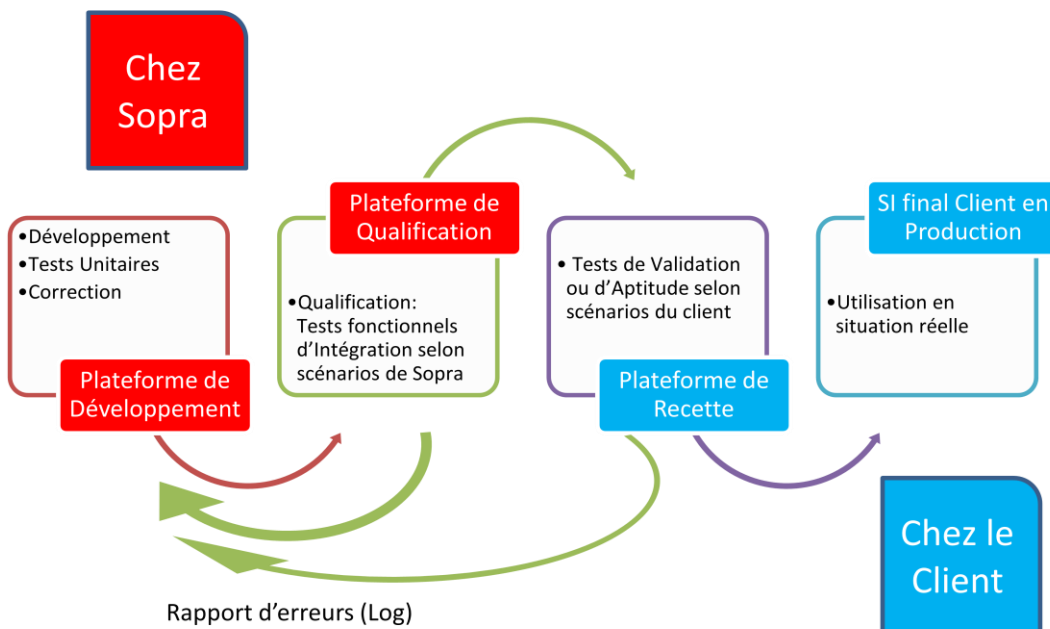
### 3.2. Les phases du projet

a. Les quatre grandes phases



#### Les grandes phases du projet

Je participe actuellement à la troisième phase du projet qui consiste en l'étude, - analyse et conception, et le développement des solutions applicatives. On y réalise également les différents types de tests selon la plateforme de travail et la maturité de la solution applicative. Cette troisième phase est plus ou moins avancée selon l'ordonnancement des modules du projet.

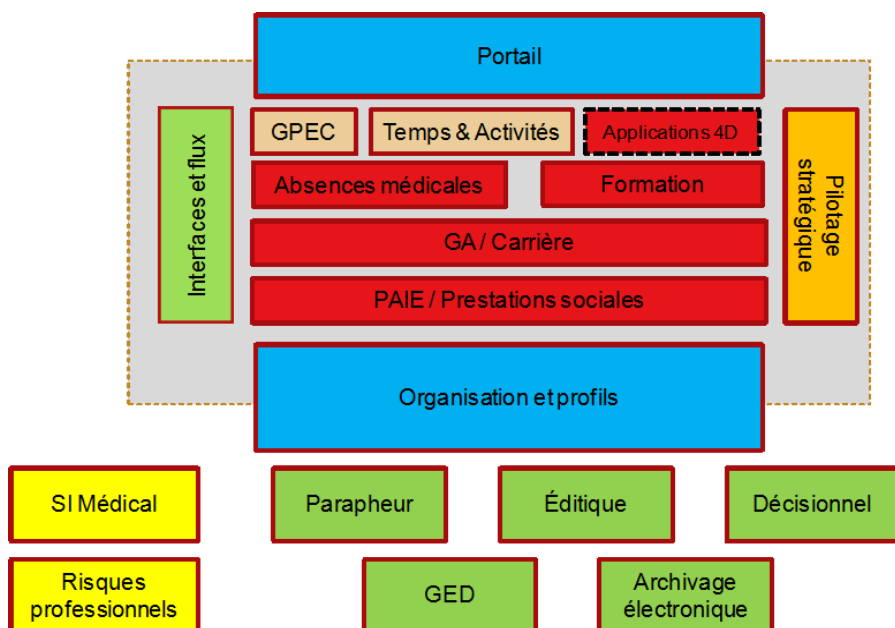


#### Cycle de Vie d'une solution applicative : développements, tests et corrections

b. L'intégration du travail dans le projet avec SVN

Le travail de gestion de version ou versioning est réalisé grâce à Tortoise SVN. Pour rappel, la gestion de version est primordiale pour la sauvegarde ainsi que pour le travail collaboratif.

3.3. Les modules du projet



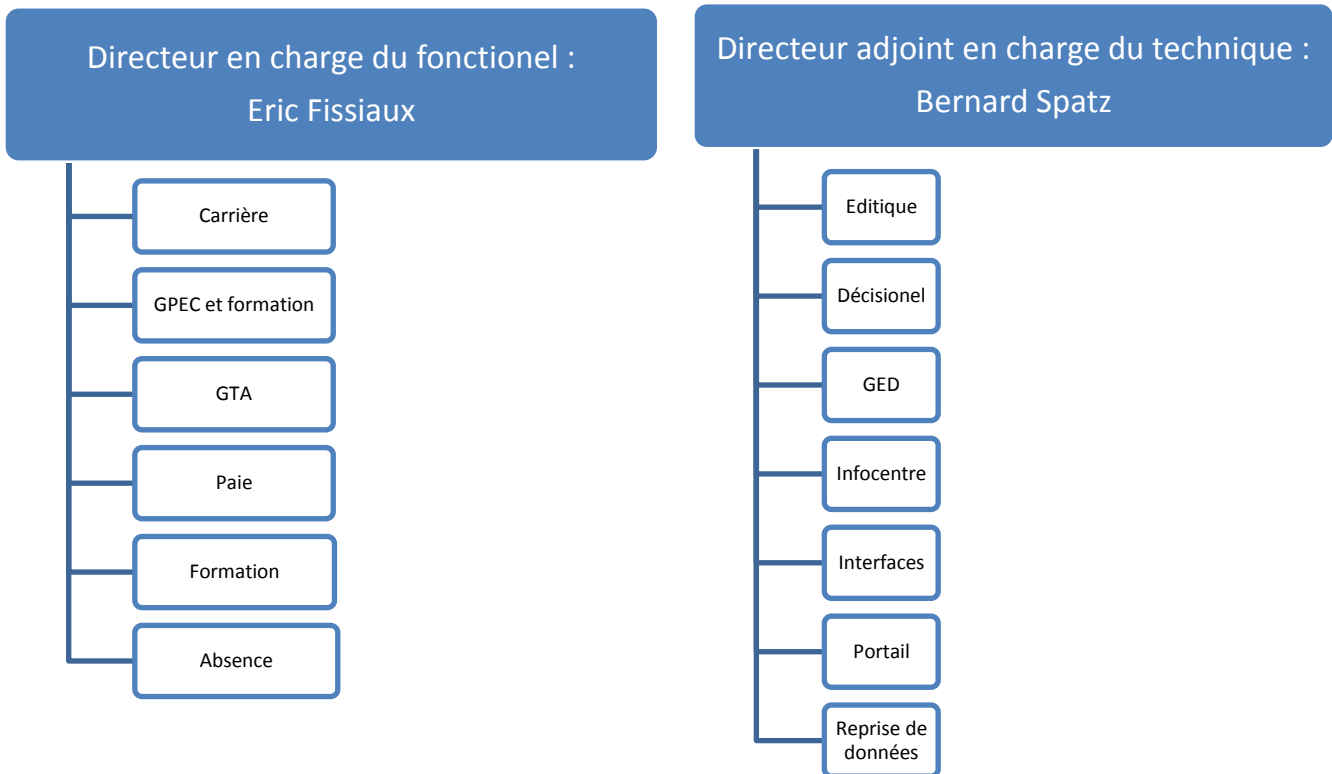
Architecture fonctionnelle du SIRH prévisionnel proposé par Sopra

Le projet de refonte du SIRH nécessite à la fois des connaissances techniques à travers les différentes technologies utilisées et à la fois des connaissances métier, à savoir des connaissances de la gestion RH du Secteur Public. Le projet est donc géré sur les deux aspects.

Le directeur en charge du fonctionnel est Eric Fissiaux. Il dirige l'équipe basée à Puteaux.

Le directeur en charge de la technique est Bernard Spatz, mon tuteur. Il dirige l'équipe basée à Aix.

Les deux équipes travaillent en collaboration sur le projet et notamment lors des ateliers de définition du besoin avec le client.



Mes possibilités d'intervention sur le projet relèvent d'abord de la technique.

Durant mon stage, je suis intervenu sur les modules Editique et Reprise de données. Mon rapport de stage en fera donc l'analyse.



## 4. Les missions confiées

---

### 4.1. Le module Editique

Le client souhaite réaliser l'édition de ses documents officiels dans le nouveau SI proposé par Sopra. L'objectif du module Editique est d'organiser dans un document à imprimer des informations d'une part selon une charte graphique officielle et donc rigoureusement définie, et d'autre part selon des règles de gestion du client. Ces informations seront stockées et gérées dans différentes bases de données selon leur nature.

La solution éditique retenue par Sopra est la suite StarJet APPIC, déjà utilisée par les équipes Pléiades à Puteaux dans d'autres projets.

Je suis intervenu pendant les deux premiers mois et demi de mon stage sur ce module à son commencement. J'ai donc participé aux différentes phases de la conception :

- Répertorisation et classification des documents mis à disposition par le client
- Factorisation de parties graphiques et textuelles entre documents de même typologie
- Conception :
  - o Du flux éditique, de la BDD Pléiades jusqu'aux traitements de l'outil éditique
  - o De l'organisation des données dans le fichier de données consommé par l'outil éditique
  - o De l'extraction des données à partir des documents officiels fournis par le client, ainsi que de la classification de ces données
- Étude de faisabilité :
  - o Développement de programmes éditiques sous StarJet Design pour mettre en œuvre des fonctionnalités du processus éditique
  - o Estimation du temps d'exécution de ces programmes

Pour consolider et partager mon travail d'étude et de conception, j'ai produit deux SFD, le premier global portant sur l'organisation du processus éditique et le second détaillant la solution technique sous l'outil d'éditique.

Par ailleurs, afin d'informatiser une partie de l'extraction des données issues de documents sources, j'ai implémenté des macros VBA qui effectuent l'extraction et le classement de ces données. Je détaillerai leurs fonctionnements par la suite. J'ai en outre accompagné la livraison de ces macros VBA d'une notice descriptive des fonctionnalités proposées.

Durant cette première phase, j'ai donc pu développer des compétences d'analyse et conception, de rédaction de spécifications fonctionnelles et techniques et de développement.

## 4.2. Le module Reprise de données

Dans son SIRH actuel, le client dispose d'un mainframe DePHI gérant un parc applicatif codé en Cobol, couvrant les données relatives aux agents actifs et aux agents inactifs. Les données sont gérées dans un SGBD DB2.

La volumétrie des données est d'environ 30 mille agents à reprendre, dont 12 milles actifs, à quoi s'ajoutent les données complémentaires des différents domaines fonctionnels.

Dans l'optique de la migration vers le nouveau SIRH proposé par Sopra, Pléiades, une reprise de données sur DB2 est nécessaire pour alimenter la nouvelle BDD. Elle se déroule en deux étapes. Il s'agit dans un premier temps de définir le périmètre des données repris ainsi que les transformations requises afin que les données transformées soient compatibles au format du SIRH cible. Il s'agit ensuite d'implémenter les outils réalisant ces extractions-transformations. La solution retenue par Sopra est l'outil d'ETL Talend.

Je suis intervenu sur ce module à la suite du module d'éditique, en participant aux phases d'implémentation des jobs Talend, de qualification avec l'outil Quality Center et de test d'intégration dans la plateforme Pléiades.

Mon travail consiste alors principalement à la conception et au développement de jobs Talend de différents niveaux de difficulté, répondant à des règles de gestion spécifiées dans des SFD et de tables de correspondance entre données source et données cible. J'ai également réalisé une spécification technique pour présenter et uniformiser notre méthodologie de conception de jobs Talend.

Enfin, j'ai réalisé des tests fonctionnels à partir de plans de test formalisant les étapes de test, correction et intégration.

Durant cette deuxième phase, j'ai donc pu également développer des compétences de conception, de développement et de rédaction.

Les deux parties suivantes détailleront le travail effectué sur chacune de ces deux parties au travers d'exemples servant de fil rouge.

## 5. Le module Editique

---

### 5.1. La problématique

La difficulté du module éditique réside en la multitude d'éditions à gérer, environ 500, ainsi que les critères de gestion de chacune, c'est-à-dire qu'une combinaison de critères permet d'indiquer quels sont les scénarios à éditer. La difficulté est d'autant plus grande que les règles de gestion des scénarios d'édition ne sont pas documentées; l'équipe de développement sur Aix ne disposant uniquement que des éditions officielles mises à disposition par le client, en format papier et électronique, et ne possédant pas les connaissances fonctionnelles requises pour gérer en autonomie les règles de gestion des scénarios d'édition.

La responsabilité de l'équipe d'Aix était au départ restreinte au périmètre de l'outil éditique, les données relatives aux scénarios d'édition devant être déjà fournies à usage de l'outil éditique APPIC. Il s'avère au cours de l'analyse conception que la prise en charge de certains aspects fonctionnels spécifiques au client incombera également à l'équipe d'Aix. Comme par exemple l'intégration des textes officiels spécifiques aux collectivités territoriales dans la base de données Pléiades.

Dans son SIRH actuel, le client possède un outil éditique développé sous 4D. Cet outil détermine le scénario à éditer à partir de critères et selon les règles de gestion du client, puis rassemble les fragments correspondant au scénario pour constituer l'édition. Cet outil répond bien aux besoins actuels du client. Cependant, il possède des inconvénients :

- Il a été conçu sur-mesure et réajusté progressivement selon l'utilisation courante du client. Du fait de la trop grande spécificité de cet outil, le module Éditique doit être totalement repensé dans la perspective d'une intégration dans le nouveau SIRH Pléiades.
- Il doit gérer une base de données de fragments d'édition ainsi que des traitements complexes pour associer ces fragments. De plus, les informations contenues dans ces fragments sont à modifier manuellement, alors qu'une gestion par BDD est préférable.
- Il ne dispose d'aucune documentation détaillée. Ainsi une étude approfondie du code des traitements est nécessaire si l'on veut en déduire les règles de gestion.

La solution éditique proposée par Sopra doit pouvoir s'intégrer dans le nouveau SIRH Pléiades tout en étant plus fonctionnelle et plus factorisable.

Pour rappel, la solution éditique réalisée par Sopra se basera sur l'outil éditique StarJet APPIC dont le module de développement est StarJet Design SJD.

Pour résoudre la problématique du grand nombre d'éditions à éditer, on classe les documents à éditer par typologie dans le but de proposer un projet éditique par typologie, le format d'édition dans chaque typologie devant être rigoureusement identique.

D'après le document de référence sur la liste d'éditions, 9 typologies d'édition ont été identifiées :

- Arrêté
- Attestation
- Contrat
- Convention
- Courrier Agent
- Courrier Externe
- Courrier Interne
- Etat de gestion individuel
- Tableau

Le type d'édition « Arrêté » est le plus volumineux avec 283 éditions, suivi des trois typologies de courrier avec 108 éditions.

Durant la première phase d'étude, nous avons mis l'accent sur le type d'édition « Arrêté », le plus important par son volume et par sa priorité. Je vais donc illustrer l'étude fournie au travers d'un fil rouge portant sur le traitement d'un Arrêté.

Concernant la problématique de la gestion du scénario à éditer, elle doit être prise en charge par les fonctionnels Pléiades en consultation des règles spécifiques du client. En effet, le standard de Pléiades permet de déterminer à partir de cinq critères au plus, quels sont les scénarios à éditer, et ce grâce à des tables d'aiguillage. Il s'agit donc d'un travail de paramétrage et de modélisation lorsque cinq critères ne suffisent pas à déterminer le scénario d'édition.

L'avantage des tables d'aiguillage par rapport à des tests conditionnels est qu'on peut factoriser les traitements. En effet, il suffit de modifier les données une seule fois sur les tables et non autant de fois que de tests conditionnels utilisant ces données.

Je n'ai pas participé à cette partie, ce rapport ne donnera que le principe de la gestion de scénarios au niveau de Pléiades.

## 5.2. Proposition théorique de résolution

La solution éditique proposée par Sopra devra permettre le déclenchement d'une édition à travers une IHM intégrée dans Pléiades. Le schéma suivant en illustre le principe.

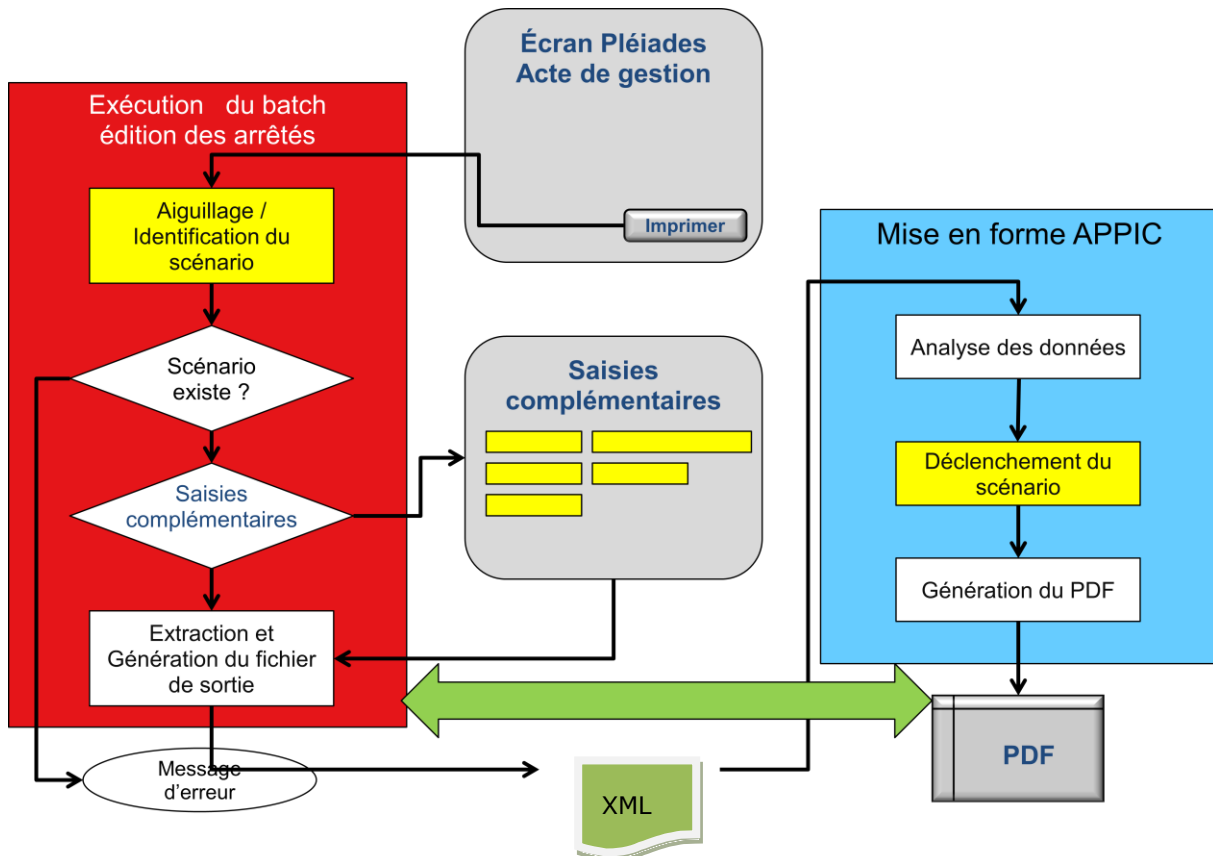
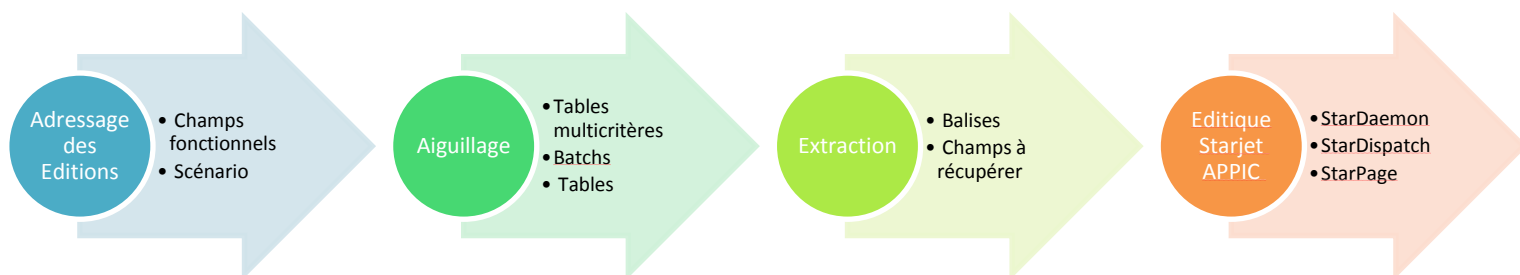


Schéma IHM du déclenchement d'édition

Au déclenchement de l'impression, un processus d'analyse des données de contexte est déclenché. Celui-ci permet d'identifier le scénario d'édition à déclencher. Certains scénarios pourront afficher une page de saisies de données complémentaires. Une fois saisie, le processus poursuit sur l'extraction des données nécessaires à l'édition du scénario. Un fichier au format XML est ainsi créé et mis à disposition de l'outil de mise en forme APPIC.

Le processus APPIC de surveillance de l'arrivée d'un fichier, déclenche les traitements de mise en forme des données à partir de masques d'édition. Un PDF est alors généré et affiché à l'utilisateur qui peut ensuite l'imprimer.

Le processus de déclenchement d'édition est mis en œuvre au travers du processus éditique global suivant :



### Schéma du processus éditique global

Ce processus se décline donc en quatre principales étapes.

#### 5.2.1. Adressage des Éditions

Il s'agit de l'identification du scénario d'édition à partir des informations renseignées par l'utilisateur dans l'écran Acte de Gestion.

Pour que cette identification soit possible, il faut pouvoir associer le scénario aux informations et aux contenus qui le constituent. Ainsi l'adressage des éditions est déterminé par trois éléments :

- La décision qui conduit à l'arrêté,
- La situation avant l'arrêté,
- La situation après l'arrêté.

En standard Pléiades, les critères pouvant être utilisés pour définir la situation avant l'arrêté et la situation après l'arrêté sont les suivants :

- Famille de Statut
- Catégorie hiérarchique
- Filière professionnelle
- Cadre emploi
- Situation de Carrière
- Type d'engagement
- Type de temps non complet
- Position administrative
- Cas Position
- Modalité de travail
- Aménagement de travail
- Catégorie d'absence
- Nature de l'absence
- Type de sanction

La détermination des champs fonctionnels à partir de la maquette d'édition permet d'identifier les critères à utiliser dans Pléiades.

### 5.2.2. Aiguillage fonctionnel

A partir des critères fournis et du motif de décision, Pléiades va procéder à l'aiguillage des tables multicritères pour en déduire les scénarios à éditer, le domaine fonctionnel ainsi que les batchs correspondants. Ces batchs vont extraire des tables de données les informations pertinentes, selon le domaine fonctionnel, comme par exemple les arrêtés de Décision d'avancement ou les arrêtés réglementant les Absences.

### 5.2.3. Extraction et constitution d'un XML de données

Les batchs Pléiades vont écrire les informations extraites des tables dans des fichiers XML. Ces informations concernent les données des agents réparties sur les trois éléments de l'adressage d'édition, les textes officiels de l'édition ainsi que des paramètres éditiques à destination de la suite StarJet APPIC.

### 5.2.4. Editique Starjet APPIC

Le fichier XML de données est alors envoyé au module éditique Starjet APPIC.

Il est dirigé vers le bon programme d'édition avec Star Daemon grâce à l'identifiant du scénario d'édition.

Ce programme est pris en charge sous StarJet Design. Le programme traitant les Arrêtés (ou assimilés) effectuera les principales étapes suivantes :

- Lecture ligne par ligne du XML
- Impression avec mis en page de la ligne lue s'il s'agit d'un texte officiel
- Renseignement des parties spécifiques à l'agent

Une fois l'édition effectuée, elle sera dirigée en impression vers l'imprimante avec retour à l'interface Pléiades dans un écran de contrôle.

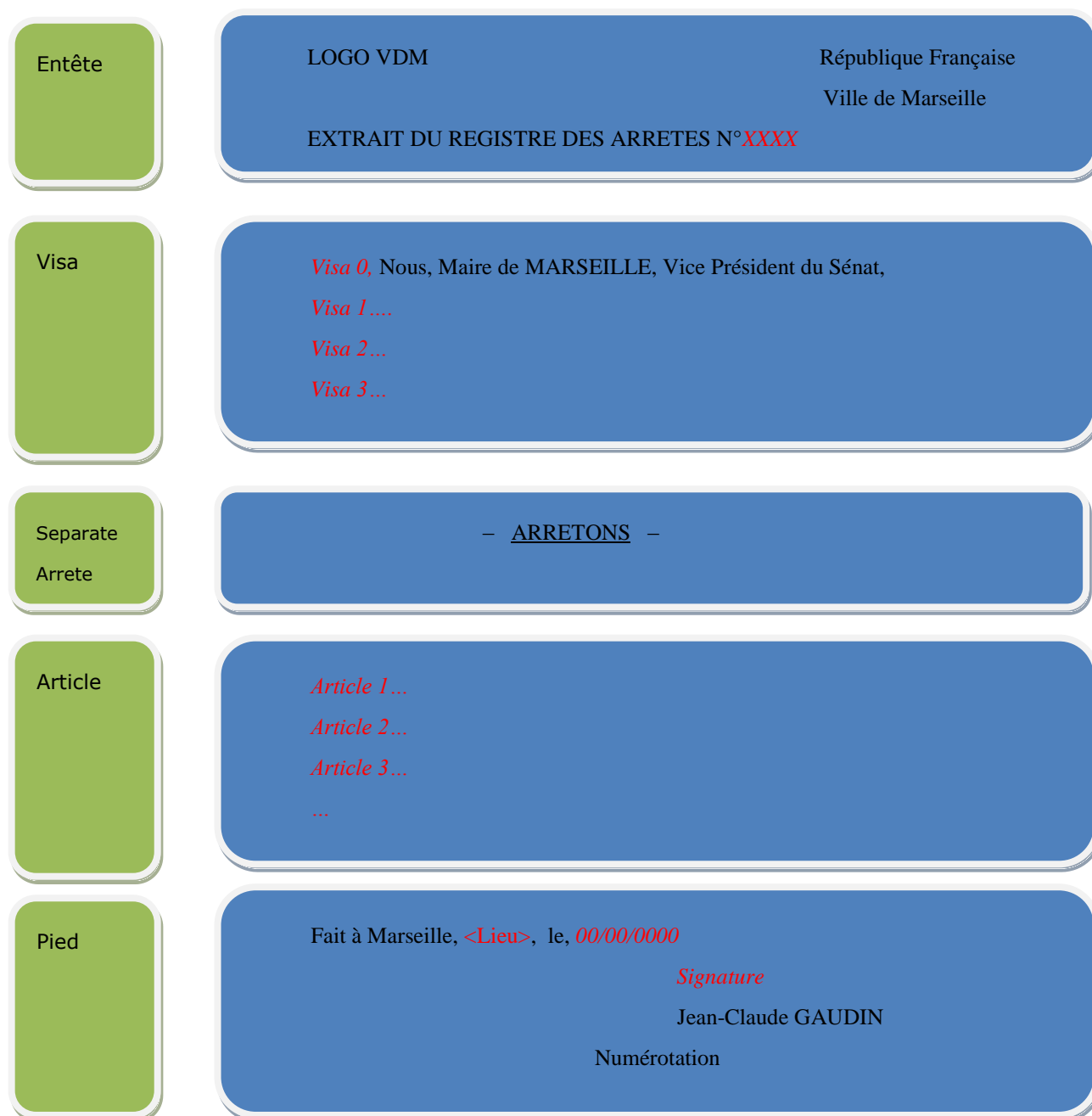
Les traitements effectués par ce module seront plus détaillés dans le paragraphe « Prise en charge côté APPIC ».

### 5.3. Mise en place de la solution

Après avoir décrit l'objet d'édition, j'expliquerai comment le décomposer de façon à stocker ses informations et ses contenus, puis enfin comment reconstruire l'objet de départ.

#### 5.3.1. Constitution d'une édition

Le fil conducteur de notre étude est l'édition d'un document de type Arrêté. Le schéma ci-dessous représente le format général d'un Arrêté :





On peut identifier cinq zones (Entête, visa, Séparateur Arrêté, Article, Pied) dont chacune est prise en charge indépendamment par un programme sous SJD.

Chacune de ces parties contient des textes officiels et des données de l'agent.

#### a. Point de départ

Nous prenons comme point de départ l'édition fournie par le client sous format papier et sous format électronique sous forme d'un document Word ou d'un document OpenOffice.

#### b. Objectif

Il s'agit de décomposer les informations graphiques et textuelles de l'édition à traiter et de constituer une BDD rassemblant ces informations.

Les informations textuelles peuvent être rassemblées dans un fichier XML. La constitution souhaitée de ce dernier sera décrit dans le paragraphe suivant.

Les informations graphiques telles que les logos, les formats de caractères et les dispositions graphiques sont gérées autrement, soit dans des bases de données de fichiers, soit au travers des programmes éditiques (cf. Overlays). Leur gestion est en cours d'étude.

## c. Constitution souhaitée du XML

Le XML attendu doit comporter une partie de variables caractérisant l'agent dans la FP, partie existante actuellement. A cela s'ajoute les parties Visas et Articles qui contiennent les textes officiels.

Une illustration est donnée ci-après :

```

<Data>
  <numPosteGst VALUE=""/>
  <RefActe VALUE="201400001"/>
  <NomDuFichier VALUE="13-123123123-20140404-AI-1-1_0"/>
  <!-- Partie existante sur l'agent -->
  <Agent>
    <MatriculeAgt VALUE="20140001"/>
    <TypeFctPublique VALUE="T"/>
    <LbQualitAgt VALUE="Monsieur"/>
    <QualiteAgt VALUE="M"/>
    <NomAgt VALUE="DETREZ PERIER"/>
    <NomUsuelAgt VALUE="DETREZ PERIER"/>
    <PrenomAgt VALUE="Bruno"/>
    <DateNaissAgt VALUE="05/05/1969"/>
    <LieuNaissAgt VALUE="Somain"/>
    <Commune VALUE="MARLY LE ROI"/>
    <Pays VALUE="France"/>
  </Agent>

  <!-- Partie à créer sur les visas -->
  <visas>
    <visa>Vu le Code Général des Collectivités Territoriales</visa>
    <visa>Vu le Code de la Sécurité Sociale, et notamment ses articles L 161-9-3, L 168-1 et suivants</visa>
    <visa>Vu la Loi N° 84/53 du 26 janvier 1984, portant dispositions statutaires relatives à la Fonction Publique Territoriale, et notamment son article 57-10°</visa>
    <visa>Vu la demande de congé de solidarité familiale, formulée par #Civilite#. #NomAgent# à compter du #DateDebutConges# pour une durée de #DureeConges#</visa>
  </visas>

  <!-- Partie à créer sur les arrêtés -->
  <articles>
    <article>
      Le congé de solidarité familiale de l'agent ci-après désigné :
      #Civilite#. #NomAgent#
      Identifiant : #IdentifiantAgent#
      Grade : #GradeAgent#
      Emploi : #EmploiAgent#
      Service : #ServiceAgent#
    </article>
    <article>#Civilite#. #NomAgent# est #autorise# à reprendre ses fonctions à l'expiration du congé de solidarité familiale.</article>
    <article>
      Monsieur le Directeur Général des Services est chargé de veiller à l'exécution du présent arrêté.
      Le délai de recours contentieux auprès du Tribunal Administratif de Marseille contre le présent arrêté est de deux mois à compter de sa notification.
    </article>
  </articles>
</Data>

```

Par ailleurs, le fichier XML doit comporter deux balises supplémentaires :

- La première identifiant la typologie du document indique à StarDaemon vers quel fichier StarPage aiguiller.
- La deuxième identifiant le scénario d'édition indique à StarPage quel(s) Layout(s) appeler, une fois l'aiguillage réalisé par StarDaemon fait.

En effet, on constituera un projet SJD sous forme d'un fichier codé sous Star Page, et ce par type de documents. De plus dans chaque projet SJD, on réalisera un Layout pour chaque groupe d'éditions similaires.

Remarque :

Ce fichier StarPage appellera à son tour le fichier StarJetDesign (.sjd) définissant l'environnement d'édition.

De plus, il y a potentiellement plusieurs Layouts à appeler pour un scénario d'édition, dans le cas où le scénario est constitué de plusieurs éditions à imprimer.

Je reviendrai sur les notions de Layout, Star Page, StarJet Design et Star Daemon.

#### d. Méthodologie sur la récupération du contenu de l'édition

Il s'agit désormais de déterminer les données à récupérer et comment effectuer cette récupération.

##### *Partie 1: Les données spécifiques à l'agent*

#### **Point de départ**

- La version papier de l'édition fournie par le client
- Deux exemplaires de fichiers XML fournis par des fonctionnels Pléiades :
  - 1 sur les arrêtés Décision
  - 1 sur les arrêtés Absence

Ce sont en fait les fichiers XML produits par des batchs Pléiades et donc contenant les variables du modèle de données Pléiades. Ils servent de référence pour transcrire le modèle de données source vers la cible.

#### **Objectif**

Établir une correspondance entre les champs relevés sur le document papier et les variables Pléiades.

#### **Critères d'évaluation**

- Pertinence
- Exhaustivité
- Unicité

### **Méthode utilisée**

Nous avons bâti notre analyse à partir des deux exemplaires XML sur les Arrêtés de Décision ou d'Absence. Pour établir la correspondance, nous avons constitué deux tableaux, « Arrêté » et « Absence » associant les champs variables identifiés manuellement sur la version papier et les variables Pléiades disponibles dans les fichiers XML respectifs. Nous avons également proposé une correspondance lorsqu'un champ identifié sur l'édition ne trouvait pas de correspondant dans le XML.

Les variables fonctionnelles Pléiades non disponibles dans le fichier de données XML doivent être intégrées dans le XML voire même dans le modèle de données Pléiades, et ce à cause des spécificités du client. Ces correspondances doivent être validées par l'équipe fonctionnelle à Puteaux.

### **Résultats obtenus**

Nous avons effectué la répertorisation des variables fonctionnelles nécessaires sur 276 éditions, comprenant l'ensemble des éditions de première priorité.

Par ailleurs, l'ensemble des éditions traitées et à traiter a été répertorié.

### **Remarque**

Au cours de la première phase d'analyse, une plateforme d'« Application » et une de « Modélisation » (une légère et une lourde) ont été mis à disposition de l'équipe d'Aix afin d'aborder le modèle de données Pléiades et sa mise en œuvre. Ces deux plateformes sont complémentaires. La première est l'interface graphique de l'utilisateur et la seconde permet de concevoir et modifier les modèles de la BDD Pléiades.

De plus une formation sur cet outil de modélisation a été dispensée à une partie de l'équipe d'Aix, début Août. Son application effective pourra s'envisager dans la deuxième phase du module Editique.

## Partie 2 : Les textes officiels

Les textes officiels comme les textes de visa ou d'article sont constitués de texte général à une population d'agents donnée et éventuellement des parties variables spécifiques à l'agent concerné.

### Point de départ

La version électronique de l'édition fournie, en format Words ou OpenOffice.

### Objectif

Établir une correspondance entre un scénario d'édition et les textes officiels de l'édition de ce scénario.

### Critères d'évaluation

- Maintenabilité aisée de la BDD de textes officiels
- Lisibilité de la BDD
- Intégrable dans le standard Pléiades

### Méthodes utilisées

Afin d'automatiser la récupération des textes officiels, j'ai implémenté trois macros VBA.

Le déclenchement des trois macros suit l'ordre suivant :



La première macro transfère les éditions dans un répertoire unique en veillant à ce que le document cible soit en formats « .docx » ou « .doc » et non endommagé. En effet, les éditions sont à l'origine réparties dans les différents dossiers fonctionnels.

Les deux macros suivantes importent des Visas et des Articles respectivement, et ce depuis les éditions centralisées dans le répertoire unique.

Elles effectuent les traitements suivants.

Les traitements bouclent sur l'ensemble des documents de format .doc\* du répertoire de copie renseigné par l'utilisateur.

Pour chaque document, la macro :

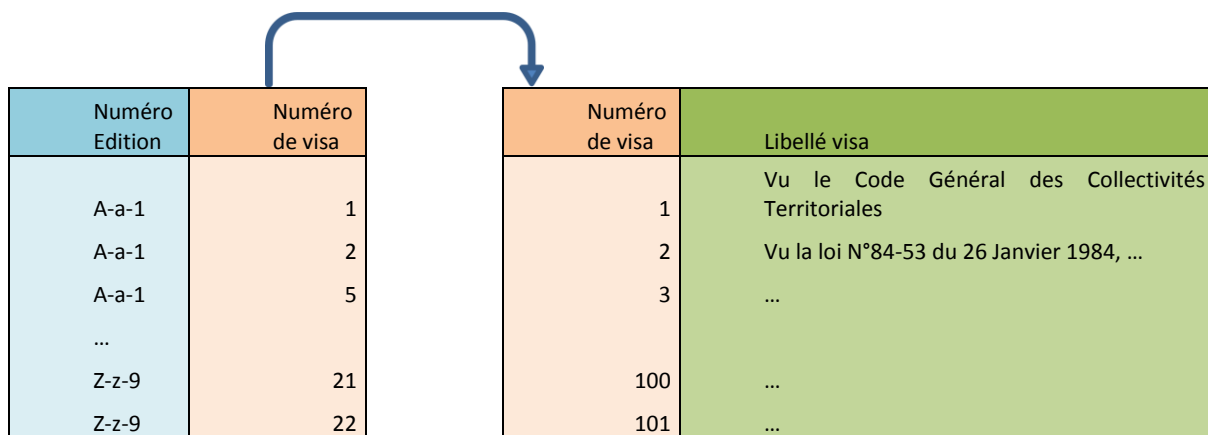
- Récupère la partie Visa (resp. la partie Article) à l'aide de délimiteurs spécifiques de début et de fin
- Importe cette partie dans une feuille Excel correspondante à l'édition en cours, cette feuille contiendra la liste des Visas (resp. des Articles) telle que chaque visa (resp. article) soit isolé dans une cellule Excel.

Chacune des macros VBA fournira également une feuille principale où sont regroupés tous les Visas (resp. Articles) de tous les éditions traitées, avec une identification de l'édition respective à laquelle se rattache le visa (resp. l'article).

Enfin, à l'aide d'une étape manuelle, chaque macro constituera deux feuilles Excel, l'une contenant une table répertoriant tous les visas (resp. articles) de façon unique et avec numérotation, l'autre listant les visas (resp. articles) de toutes les éditions traitées et faisant la correspondance entre le numéro du scénario d'édition et le numéro du visa (resp. article).

**Résultats obtenus**

- o Pour chacune des parties Article ou Visa, on obtient deux tables de correspondance :
  - Numéro scénario/ numéro de visa (resp. article)
  - Numéro de visa (resp. article)/ libellé du visa (resp. article)



Ces tables seront intégrées dans le modèle de données Pléiades et leurs informations seront récupérées dans le XML grâce aux batchs Pléiades.

*Partie 3 : Les informations graphiques : Logo, En tête, Pied Page, Signature*

Ces informations graphiques ont été répertoriées et factorisées à travers les différentes éditions. Une uniformisation des différentes chartes graphiques est en étude.

Ces informations graphiques doivent être stockées dans une base de données d'images exploitables par le logiciel éditique SJD.

Enfin, la mise en forme suivant les chartes graphiques du client est gérée directement dans les programmes de la solution SJD.

Nous avons ainsi décomposé une édition et stocké les informations décomposées dans différentes BDD selon la pertinence. Il reste maintenant à savoir comment récupérer ces informations une fois l'édition identifiée.

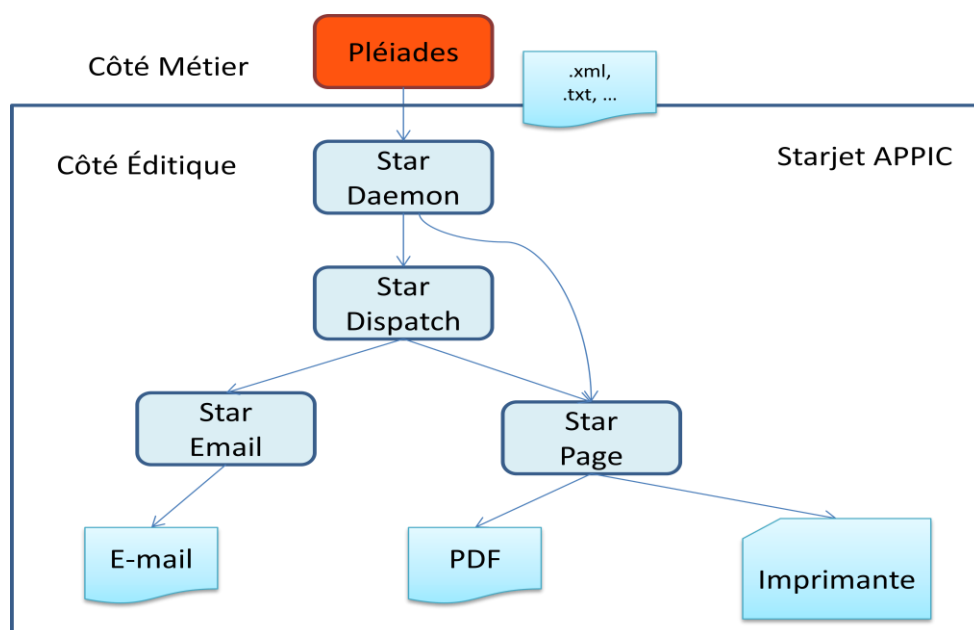
### 5.3.2. Prise en charge côté APPIC

#### a. Schémas de principe général

##### Organisation de la solution StarJet APPIC avec ses différents modules

On suppose donc que l'édition à éditer est connue et que Pléiades a produit le fichier de données XML. Ce dernier constitue le point d'entrée du périmètre éditique pris en charge par la suite StarJet APPIC.

Le schéma ci-dessous présente l'enchaînement des modules éditique.



Star Daemon est une routine qui scrute le répertoire en entrée pour détecter le fichier de donnée XML. Il lance le cas échéant le bon programme SJD écrit sous la solution StarPage.

StarDispatch permet ensuite, d'une part de rassembler les informations issues de différents fichiers de données source et d'autre part de diriger le flux de données vers différents traitements. Nous utiliserons plutôt la seconde fonctionnalité, par exemple dans le cas où l'utilisateur souhaite appliquer à la fois un traitement de mise en page d'e-mail avec Star Email et à la fois un traitement de mise en page de document officiel sous SJD. Je n'insisterai pas plus sur le module StarDispatch.

Enfin, Star Page est la solution de programmation de SJD. On reviendra plus en détail sur les fonctionnalités utilisées du module SJD.



b. Gestion des flux d'entrée : StarDaemon

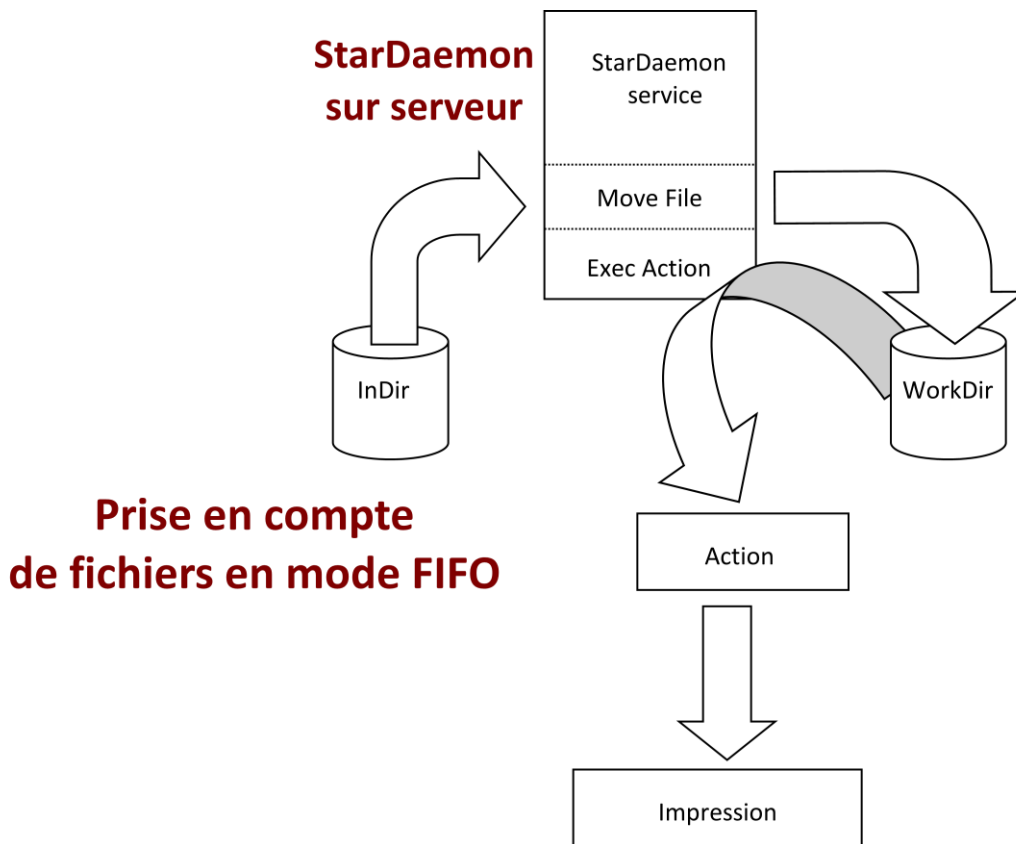


Schéma de la gestion des flux par StarDaemon

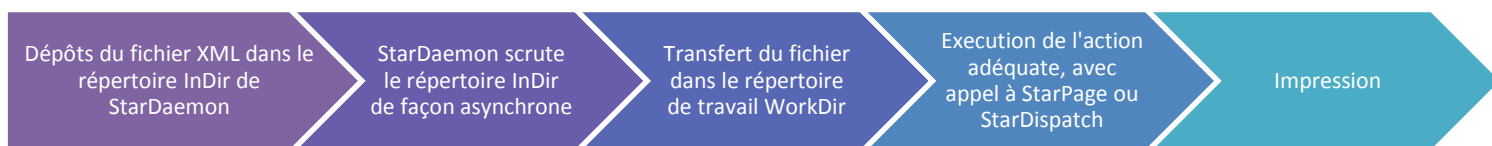


Schéma du procédé de gestion des fichiers entrants

Le rôle principal de Star Daemon est d’aiguiller le fichier XML vers le bon projet éditique mis en œuvre par un fichier Star Page. L’aiguillage est faite grâce à une balise du XML.

Une fois l’exécution du fichier Star Page faite, Star Daemon déclenchera l’impression soit vers une imprimante, soit en pdf, selon le paramétrage.

### c. Traitement des éditions : StarJet Design

#### *Notions propres à la solution SJD*

Il y a 3 notions importantes dans un projet éditique StarJet Design :

- **L'Overlay :**

Il s'agit d'un calque dont les dispositions et les contenus sont fixes. Ils sont utiles en présence de logos par exemple.

- **Les Macros :**

Il existe deux types de macro dans les projets StarJet, des macros imprimables et des macros non imprimables. Dans notre cas, nous n'utiliserons que les macros imprimables.

La macro imprimable est une procédure pouvant être appelée depuis une autre macro imprimable ou par des Layouts. Les macros gèrent et traitent des données.

Elles peuvent par exemple gérer dynamiquement des champs informationnels de taille variable, ainsi que leurs positionnements.

Elles peuvent aussi extraire des données depuis le fichier XML et les affecter aux variables internes d'un projet.

- **Le Layout :**

Il gère notamment l'ordre et les conditions d'appel des macros imprimables pour la mise en forme et le positionnement sur la maquette d'édition. Il peut également appeler des overlays.

Par ailleurs, un projet éditique StarJet Design permet de gérer ces trois types de composants par interface graphique ainsi que le développement du fichier de code Star Page ainsi qu'un debugger.

#### *Organisation d'un projet SJD*

Dans un but de factorisation des traitements des documents, nous avons opté pour deux regroupements successifs, par type puis par typologies de documents :

1. Nous créerons un projet SJD par type de document (Arrêtés, Courriers, Attestation, etc.)
2. Au sein d'un projet SJD, on pourra réaliser plusieurs Layouts correspondants chacun à un regroupement par typologie.

Par exemple, on peut traiter les arrêtés et les contrats dans un même projet Starjet Design mais avec deux Layouts. De même pour les courriers Agent, courriers Externe et courriers Interne avec trois Layouts.

### Les étapes du traitement

Un projet StarJet est associé à un fichier de données unique. Il peut prendre en entrée comme fichier de données des fichiers de type ASCII. Dans notre cas précis, nous utiliserons des fichiers XML comme source de données.

Afin de compléter une information ou dans le cas d'un besoin de combinaison de données provenant de sources différentes, StarJet offre la possibilité de pouvoir accéder à des fichiers supplémentaires grâce à la fonction externe (Extern\_fonc) : Option\_Query. Cette fonction permet d'accéder à tout type de fichier ASCII et d'en extraire les informations nécessaires.

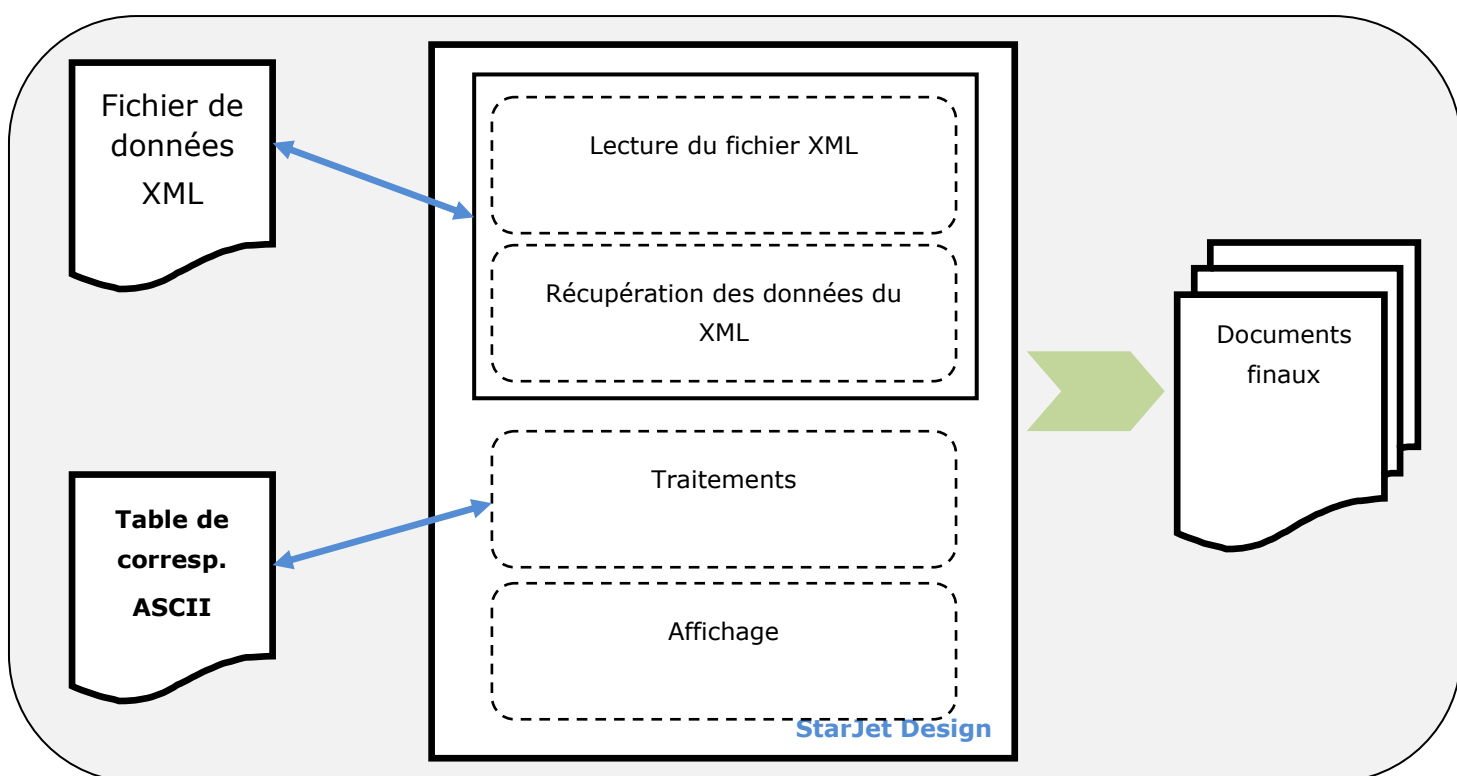
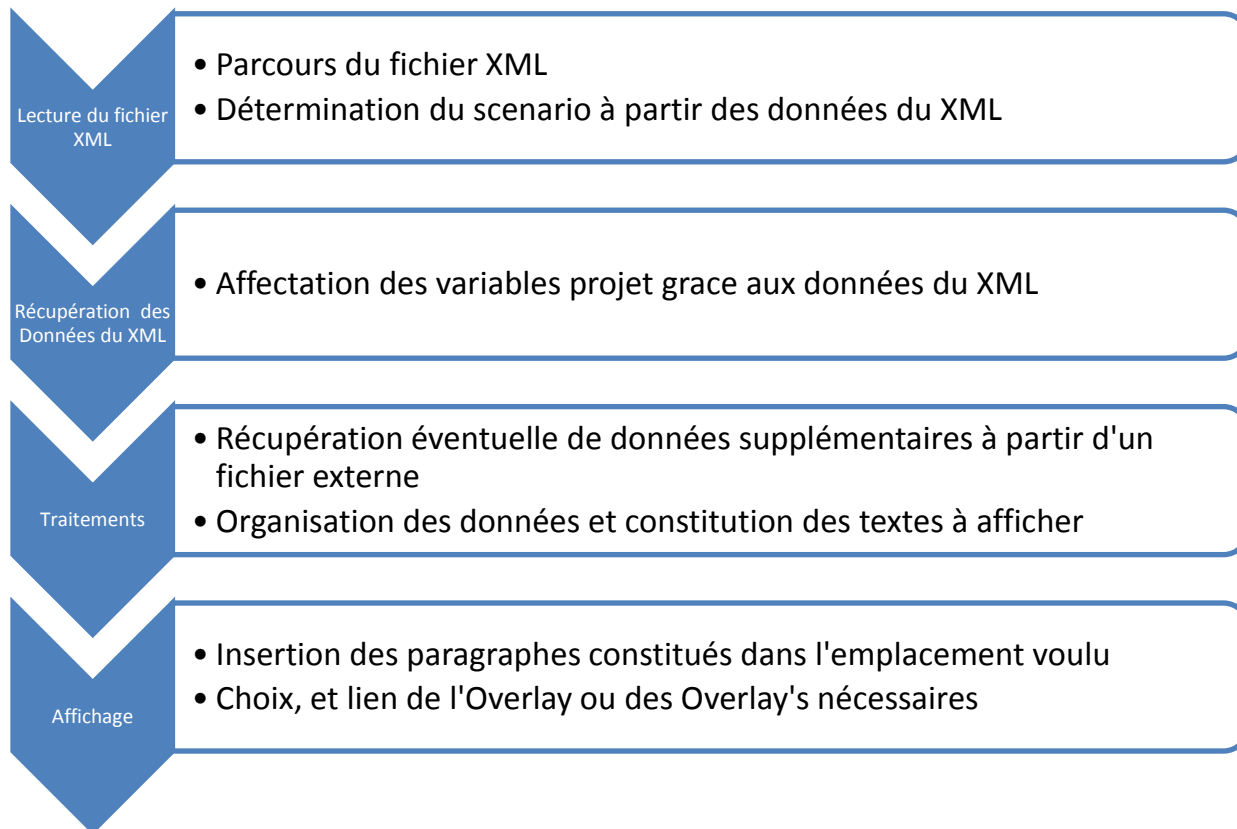


Schéma des quatre principales étapes du traitement d'un projet SJD



Processus de traitement d'un projet SJD

Le traitement principal pour une édition de typologie Arrêté boucle sur les visas puis sur les articles de l'édition en effectuant les opérations suivantes :

- Lecture du texte officiel
- Remplacement des variables par les données spécifiques de l'agent
- Impression du texte officiel avec les champs renseignés

## 6. Le module Reprise de données

---

### 6.1. La problématique

Le travail de la reprise de données est de définir des entités Pléiades sous forme de fichier texte, dans le but de leur intégration dans la BDD Pléiades; le schéma et les données provenant du modèle de données du client.

La difficulté de la reprise de données réside d'une part dans le volume de données à reprendre et d'autre part dans la complexité des règles de gestion du client qui impactent sur les traitements à réaliser.

Les données volumineuses augmentent le taux d'erreurs sur les données. Il faut donc prévoir une gestion des erreurs à trois niveaux :

- La cohérence des données source du client avec les spécifications
- La possibilité de traitement des données source par Talend
- La possibilité de traitement des données transformées en intégration de Pléiades

Remarquons au passage que l'outil Talend est adapté pour le niveau de volume repris, ce qui permet d'assurer la performance des sessions ETL lors des exécutions de jobs Talend.

Les règles de gestion du client ont déjà été retranscrites et sont données comme point de départ de mon travail. Certaines peuvent être difficiles à implémenter.

Par ailleurs, les données sont catégorisées par domaines fonctionnels, à l'intérieur desquels sont définies des entités. Ces derniers sont des tables de la base de données Pléiades. En effet, dans la plupart des cas, il existe déjà une table Pléiades pour chaque table à reprendre du client, la table Pléiades comprenant les colonnes de la table du client. Dans certaines situations, plusieurs tables du client définissent une unique table Pléiades.

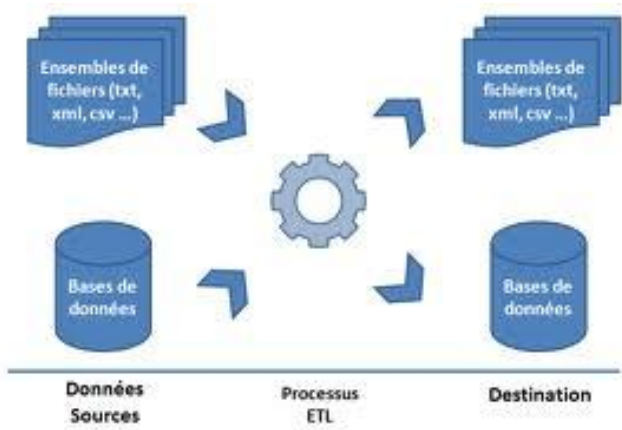
De plus, la reprise de données doit suivre un ordre de reprise des domaines fonctionnels. L'ordre de reprise, correspondant à l'ordre d'intégration dans Pléiades, n'est cependant pas détaillé dans ce rapport.

La reprise de données est mise en œuvre grâce à des jobs Talend. La qualité de leur implémentation est qualifiée par un outil de test, Quality Center.

Les parties suivantes décrivent la mise en œuvre des jobs Talend, de l'extraction-transformation des données source à l'intégration dans la BDD cible ainsi que la qualification du travail.

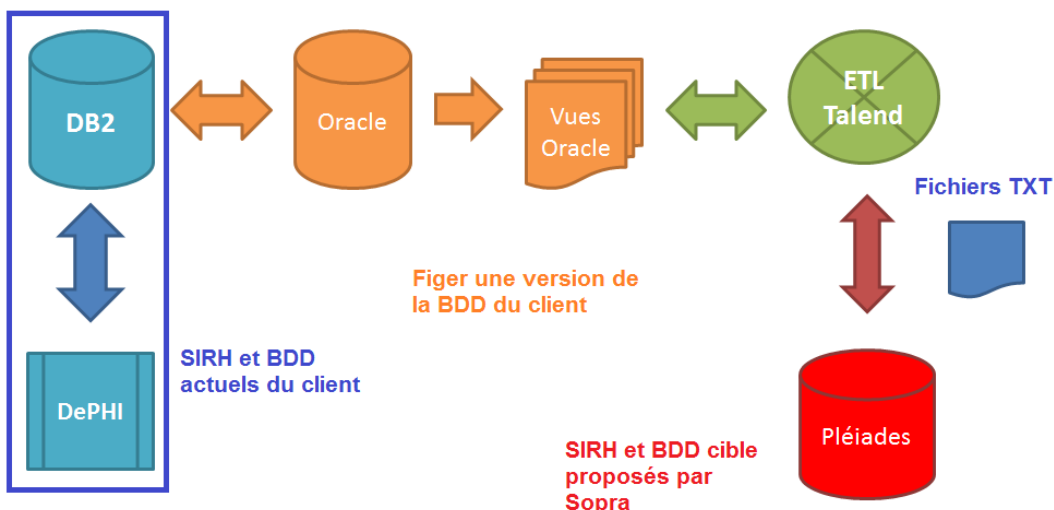
## 6.2. Proposition théorique de résolution

### 6.2.1. Schéma global : BDD source – ETL Talend – BDD cible



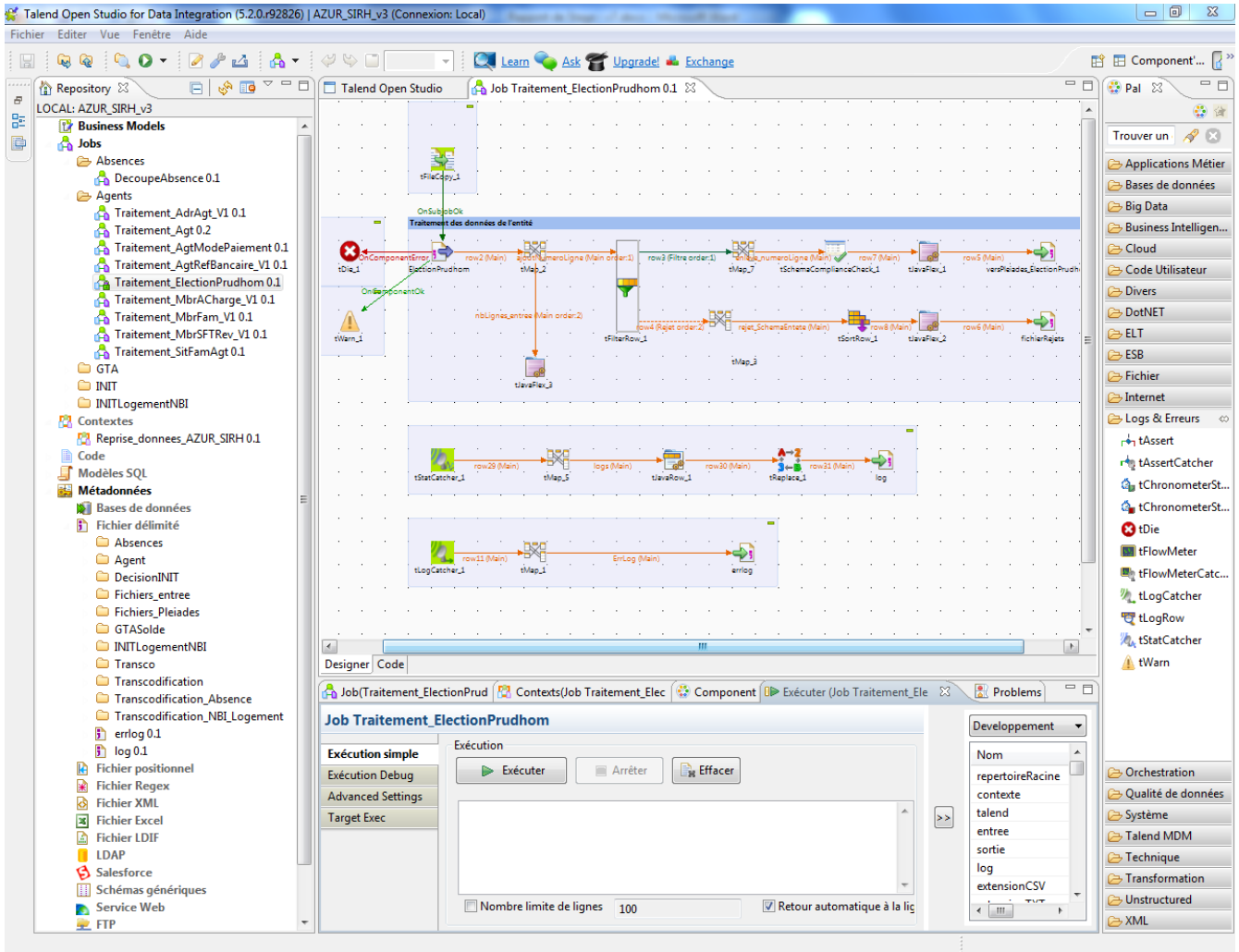
#### Processus ETL général

Le processus ETL extrait les données depuis la BDD source, les transforme et les charge dans la BDD cible. Les données sont extraites de la BDD DB2 de DePHI. Elles sont d'abord figées dans des vues Oracle. En effet, le client va continuer à exploiter son SIRH et les données de sa BDD DB2 continueront à fluctuer. Elles sont ensuite écrites dans un fichier CSV qui sera traité dans les opérations de la phase transformation par l'outil d'ETL Talend. Les données transformées alimenteront la BDD Pléiades par l'intermédiaire d'un fichier texte TXT. L'objectif étant de rendre tout le processus autonome à une intervention manuelle.



#### Processus ETL implémenté pour la partie Reprise de données

L'interface graphique Talend est présentée ci-dessous.



Au centre, on retrouve un espace de travail où l'on peut manipuler les composants Talend par glisser-déposer.

La fenêtre à gauche permet de définir des jobs, des contextes et des métadonnées. Un job est un traitement exécutable Talend. Un contexte permet de définir des variables de contexte; elles variabilisent les noms des fichiers et les chemins de répertoires. Une métadonnée définit le schéma de données d'une entité.

La fenêtre à droite propose les différents composants Talend.

Enfin la fenêtre en bas au centre permet de contrôler les paramètres d'exécution.

## 6.2.2. Reprendre chaque entité par un job Talend

Le principe de notre méthodologie de reprise de données est d'effectuer un traitement ETL pour chaque entité Pléiades à l'aide d'un job Talend.

Les jobs Talend traitant une entité doivent être indépendants les uns des autres.

## 6.2.3. Principe de déclenchement des jobs

Pour les tests unitaires, il suffit de lancer l'exécution du job Talend courant directement dans l'interface graphique.

Pour les tests d'intégration en phase de qualification, il est possible d'exporter le job Talend en exécutable. Il sera alors exporté avec trois exécutables : un .jar, un .bat et un .sh .

Sous Windows, on exécute le job Talend avec le .bat . Il faut auparavant définir les variables environnementales de Java, - CLASSPATH, PATH et JAVA\_HOME.

Un membre de l'équipe d'Aix a réalisé une macro VBA pour déclencher graphiquement les exécutions des jobs dans le but de centraliser les boutons d'appel et de contrôler en direct les résultats d'exécution.

## 6.2.4. Attendus supplémentaires : Journaux de rejets, de statistiques et d'erreurs d'exécution

En complément du fichier de résultat des transformations pour une intégration dans Pléiades, nous livrons également pour chaque entité :

- Un fichier de rejets constitué des enregistrements ne répondant pas à une ou plusieurs spécifications permettant une intégration dans Pléiades. Il répertorie donc les erreurs sur les données.
- Un fichier de statistique renseignant notamment le début et la fin des exécutions, le nombre total de lignes en entrée, le nombre de lignes correctes traitées, le nombre de ligne en rejet ainsi que le taux d'enregistrements corrects traités.
- Un fichier d'erreur qui répertorie les erreurs d'exécution d'un job Talend.



### 6.3. Mise en place de la solution

#### 6.3.1. L'organisation d'un job Talend

Les jobs traitant les entités reposent dans le cas général sur un même schéma d'organisation se décomposant en quatre sous-jobs ayant un rôle spécifique :

Sous job n°1 : Initialisation du fichier de rejets

Sous job n° 2 : Processus de transformation

Sous job n° 3 : Tri du fichier de rejets

Sous job n° 4 : Gestion du journal des erreurs et des statistiques

Libellé du sous-job	Traitements effectués
Initialisation du fichier de rejets	Placer le nom des colonnes comme première ligne pour obtenir l'entête du fichier de rejets.
Processus de transformation	Traiter des données de l'entité : <ul style="list-style-type: none"> <li>- Gestion des Erreurs :                             <ul style="list-style-type: none"> <li>o sur le format de l'entête des enregistrements</li> <li>o sur le format des dates</li> </ul> </li> <li>- Gestion des erreurs de transcodification</li> <li>- Transcodification dans les cas possibles</li> <li>- Troncature des champs excédents la taille maximale</li> <li>- Mettre en œuvre les règles de gestion du client</li> </ul>

Tri du fichier de rejets	Trier les lignes du fichier de rejets. Le traitement de ce sous-job peut être incorporé dans le sous job n°2 dans les cas simples.
Gestion du journal des erreurs et des statistiques	Récupérer dans des fichiers dédiés les informations relatives aux statistiques et aux erreurs dues à l'exécution du job.

### 6.3.2. La définition du schéma de données par des méta-données

Cette étape importante permet d'initialiser le schéma des entités à l'aide du fichier de données en entrée.

**File - Step 3 of 3**  
Update an existing Metadata File on repository  
Define the setting of the parse job

**Paramètres de fichier**

Encoding: windows-1252

Séparateur de champs: Semicolon | Caractère correspondant: ";"

Séparateur de lignes: Standard EOL | Caractère correspondant: "\n"

**Paramètres du caractère d'échappement**

CSV |  Délimité

Caractère d'échappement: Vide

Entourage du texte: Vide

Découper la ligne avant le champ

**Lignes à ignorer**

Si des lignes doivent être ignorées, spécifiez les paramètres suivants

En-tête:  1

Pied de page:

Ignorer les lignes vides

**Limite de lignes**

Si le nombre de lignes doit être limité, spécifiez ce nombre.

Limite:

Preview | Sortie

Définir les lignes d'en-tête comme nom de colonnes | Rafraîchir l'aperçu

Code de l'entité	Type de mise à jour	Nom de l'utilisateur	Agt.matricule	BgnDate	EndDate	NatureAbsence.code	TypeAbs.code	duree	UniteD
DecoupeAbsenc01	a	DecoupeAbsenc09		20010615	20090308	A41			
DecoupeAbsenc02	a	DecoupeAbsenc10		20010306	20090309	AT			
DecoupeAbsenc03	a	DecoupeAbsenc11		20060311	20090310	GM			
DecoupeAbsenc04	a	DecoupeAbsenc12		20070912	20090311	CLD			
DecoupeAbsenc05	a	DecoupeAbsenc13		20110207	20120312	CLM			

Exporter en tant que contexte | Revenir au contexte précédent

< Retour | Suivant > | Terminer | Annuler

En effet, dans un premier temps nous n'avons pas accès aux BDD du client. Nous constituons nous même des fichiers de tests unitaires à partir des spécifications des modèles de données. Ces fichiers servent donc de point d'entrée des traitements ETL Talend.

Lorsque nous aurons accès aux BDD du client, il suffit de définir une connexion à ces BDD tout en précisant les métadonnées.

Cette étape permet ensuite de définir le format et la longueur des champs en accord avec les spécifications sur les champs des entités Pléiades.

Description du schéma

Column	Clé	Type	<input checked="" type="checkbox"/>	N..	Modèle de date (Ctrl...	Longueur	Precision	Aucune	Comment...
Code_de_l_entite	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			15	0		
Type_de_mise_a_jour	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			1	0		
Nom_de_l_utilisateur	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			15	0		
Agt_matricule	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			8	0		
BgnDate	<input type="checkbox"/>	Date	<input checked="" type="checkbox"/>		"yyyyMMdd"	8	0		
EndDate	<input type="checkbox"/>	Date	<input checked="" type="checkbox"/>		"yyyyMMdd"	8	0		
NatureAbsence_code	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			3	0		
TypeAbs_code	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			1	0		
duree	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			4	0		
UniteDureeAbs_code	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			1	0		
dtBen	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			10	0		

### 6.3.3. Les traitements spécifiques

Talend fournit une multitude de composants qui permettent chacun de réaliser un traitement spécifique. Ceux qui nous intéressent sont les suivants, car ce sont les plus fréquents :

- Transcodifications (jointure par tMap) :  
Conversion d'une donnée d'entrée en une donnée cible grâce à une table de référence.
- Règles de gestion (tJavaRow) :  
Le tJavaRow permet d'écrire du code Java afin d'effectuer des opérations non standard de Talend.  
Il réalise les mêmes calculs pour toutes les lignes en flux d'entrée.
- Filtre sur les données (tFilterRow) :  
Filtre les lignes répondant à des critères sur les colonnes.
- Contrôle de la qualité des données (tSchemaComplianceCheck)
- Démultiplication d'enregistrements (tFlowTolterate et tFixedFlowInput)
- Contrôle des flux : compter le nombre de lignes (tJavaFlex) :  
Le tJavaFlex permet aussi d'écrire du code Java. Il permet cependant d'appliquer des calculs d'entrée et de sortie de flux tout en gérant des calculs communs à toutes les lignes en flux d'entrée.

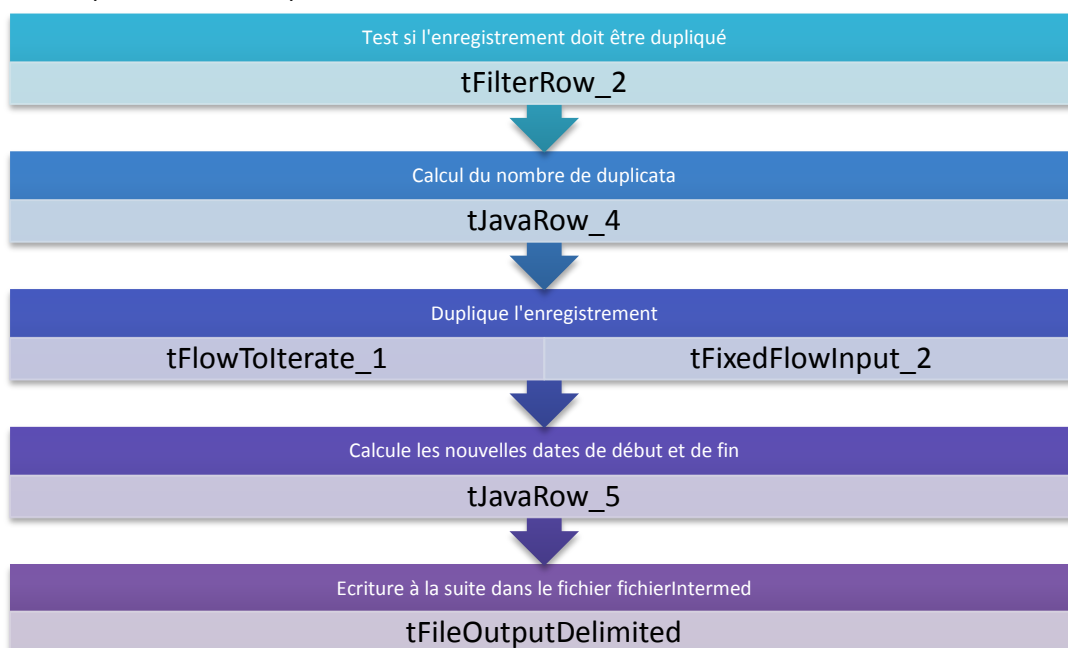
L'utilisation de ces principaux traitements est illustrée à travers une étude de cas, celui de l'entité DecoupeAbsence.

Traitement 1 : Duplication des enregistrements sous condition.

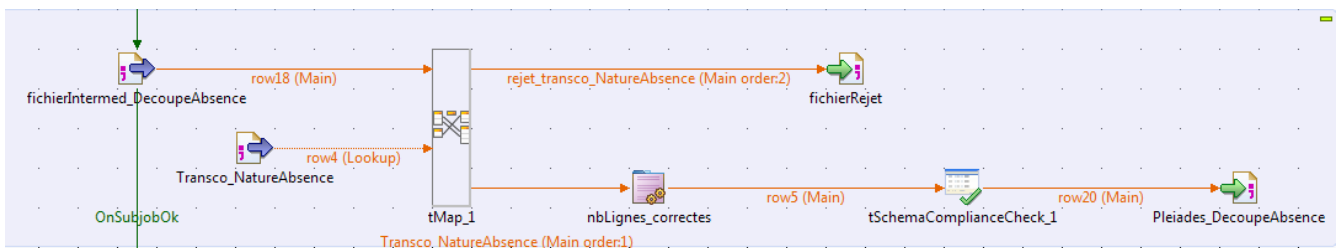
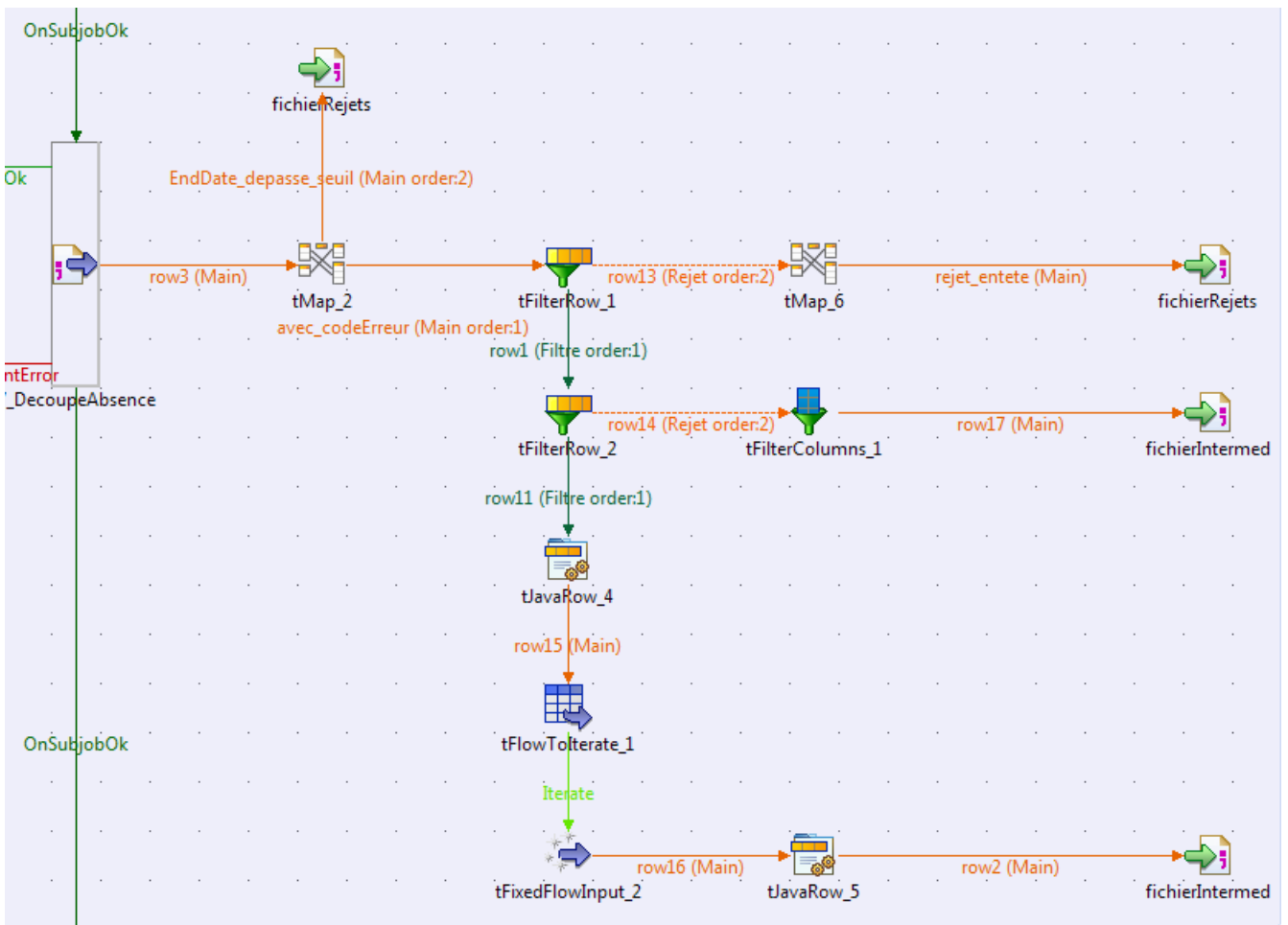
La principale difficulté de l'entité DecoupeAbsence est sa règle de duplication des enregistrements.

Elle porte sur les colonnes de début et de fin d'une absence. Il faut que pour chaque enregistrement en intégration de Pléiades, la durée d'une absence soit inférieure à 999 jours. Or certaines absences provenant des données du client dépasse ces 999 jours. Il faut alors dupliquer cette absence autant de fois qu'il y a d'intervalles de 999 jours plus un si nécessaire. Les dates de début et de fin des duplicata doivent être recalculées, les autres colonnes restent inchangées.

Le schéma du processus de duplication est donné ci-dessous.



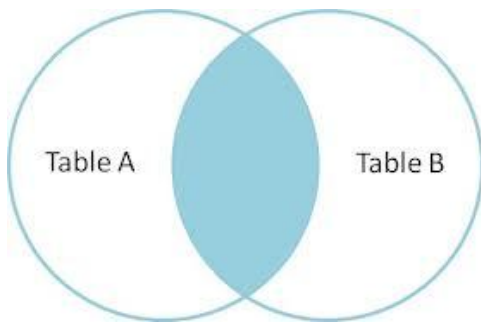
La capture d'écran de l'implémentation de cette règle de gestion est donnée ci-dessous.



**Traitement 2 : Transcodification des données source.**

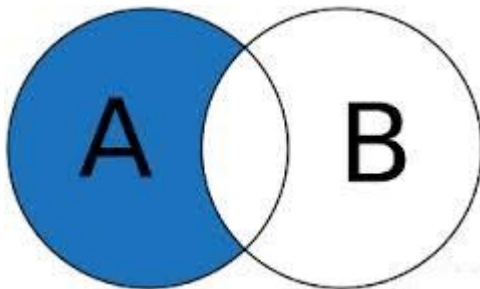
La transcodification des données source consiste à convertir des données du client en données acceptées par Pleiades, et ce grâce à une table de référence faisant la correspondance entre données source et données cible. Il s'agit d'une opération de jointure interne réalisée grâce au composant tMap\_1 sur la colonne NatureAbsence de l'entité.

Les enregistrements pouvant être transcodés peuvent être modélisés comme étant dans l'intersection en bleu clair dans le schéma ci-dessous.



Les données source qui n'ont pas de correspondance dans la table de référence sont dirigées dans le fichier des rejets. Il s'agit de la différence entre la jointure externe gauche et la jointure interne, autrement dit, il s'agit de la partie colorée en bleu foncé.

Dans la jointure externe gauche, la table de gauche est la table principale `DecoupeAbsence` et celle de droite est la table de référence.



Les données correctement transcodées sont dirigées dans le fichier d'intégration de Pléiades.

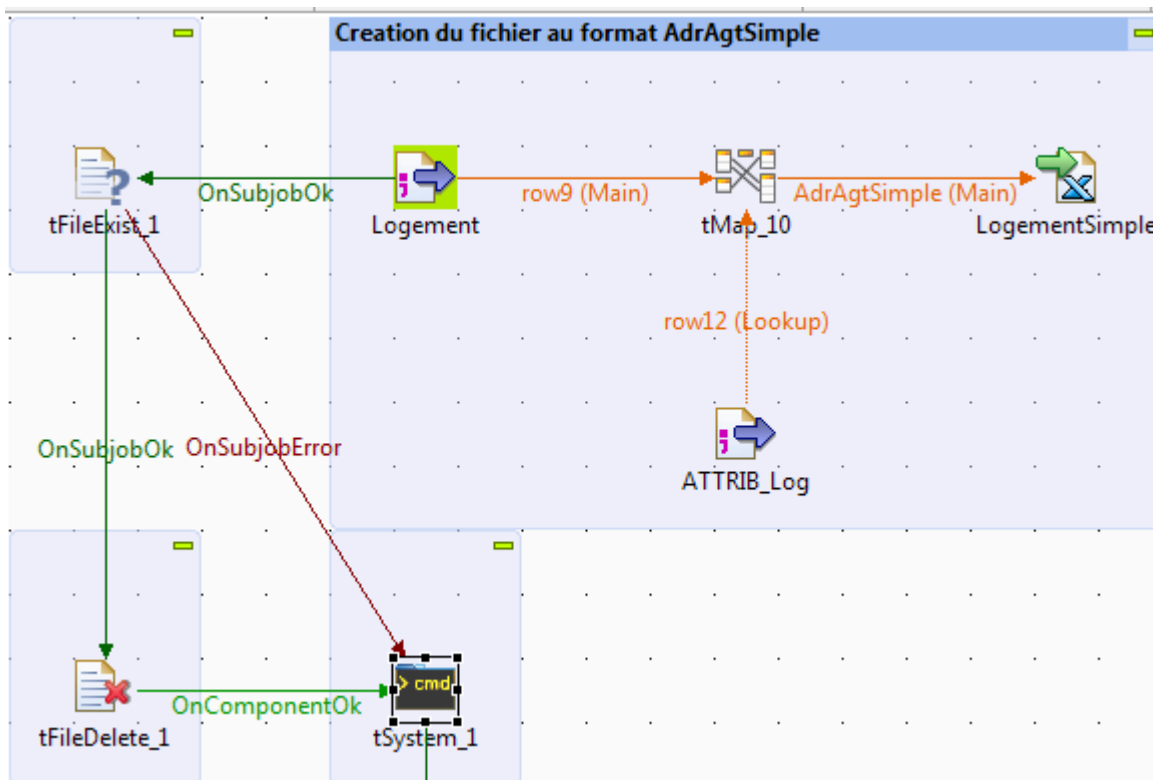
On en profite au passage pour compter le nombre de lignes de ce flux de sortie grâce à un `tJavaFlex` et on tronque les champs de longueurs excédentaires grâce à un `tSchémaComplianceCheck`.

### 6.3.4. L'intégration d'exécutables complémentaires

Il est également possible de lancer un exécutable au cours de l'exécution d'un job Talend.

On rencontre cette situation dans le cas des entités AdrAgt et Logement où l'on a besoin de lancer une macro VBA de normalisation des adresses.

Le composant tSystem permet de lancer une ligne de commande pour lancer l'exécution de cette macro.



### 6.3.5. La variabilisation des chemins et noms de fichiers par des variables de contexte

Grâce à cette étape, il suffit désormais de changer une seule fois les variables de contexte lorsque l'on change de plateformes de travail ou de norme de nommage.

Il y a en tout quatre plateformes de travail distincts, qui correspondent aux quatre phases du livrable :

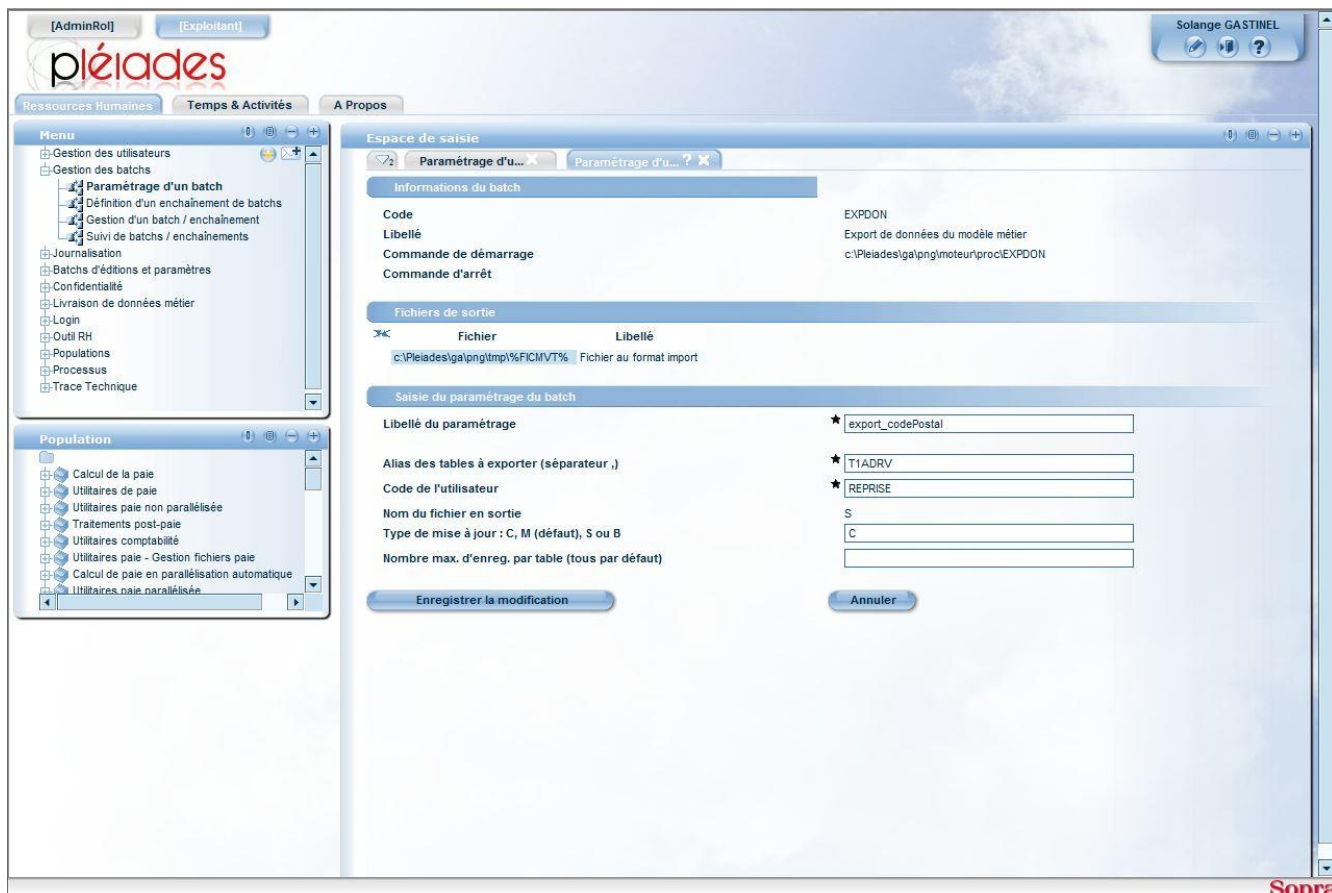
- Développement
- Qualification
- Recette
- Utilisation
- Production

La variabilisation du nommage est également utile en cas de modification de norme de nommage de répertoires ou de fichiers.

### 6.3.6. Intégration dans Pléiades

Les jobs Talend produisent des fichiers TXT destinés à être intégrés dans Pléiades. Chaque enregistrement d'un fichier TXT correspondant à une entité obéit à un schéma défini par le modèle de données Pléiades.

Pour déclencher l'import dans Pléiades, on lance le batch d'import « IMPDON » en lui précisant l'alias de l'entité Pléiades. On précise également le mode de mise à jour, à savoir en Création, en Modification, en Création/Modification ou en Suppression.



L'import dans Pléiades est réalisé au niveau du portail web, aussi appelée l'Application.



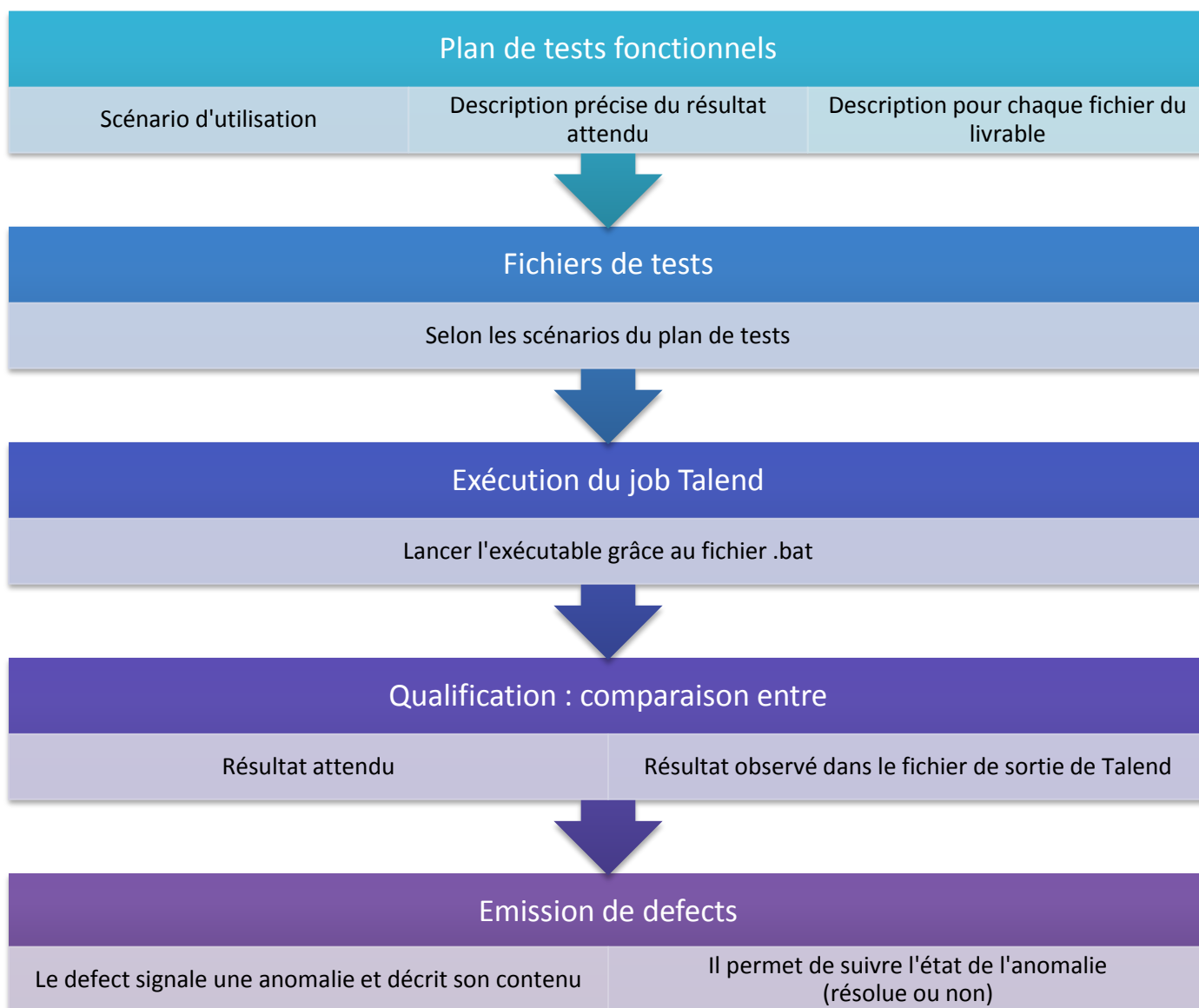
## 6.4. Qualification du travail

### 6.4.1. L'outil : Quality Center HP

Quality Center permet de formaliser les étapes de qualification dans une procédure de test.

La qualification est une étape de gestion de la qualité du développement qui consiste à vérifier si le résultat obtenu est conforme au résultat attendu spécifié.

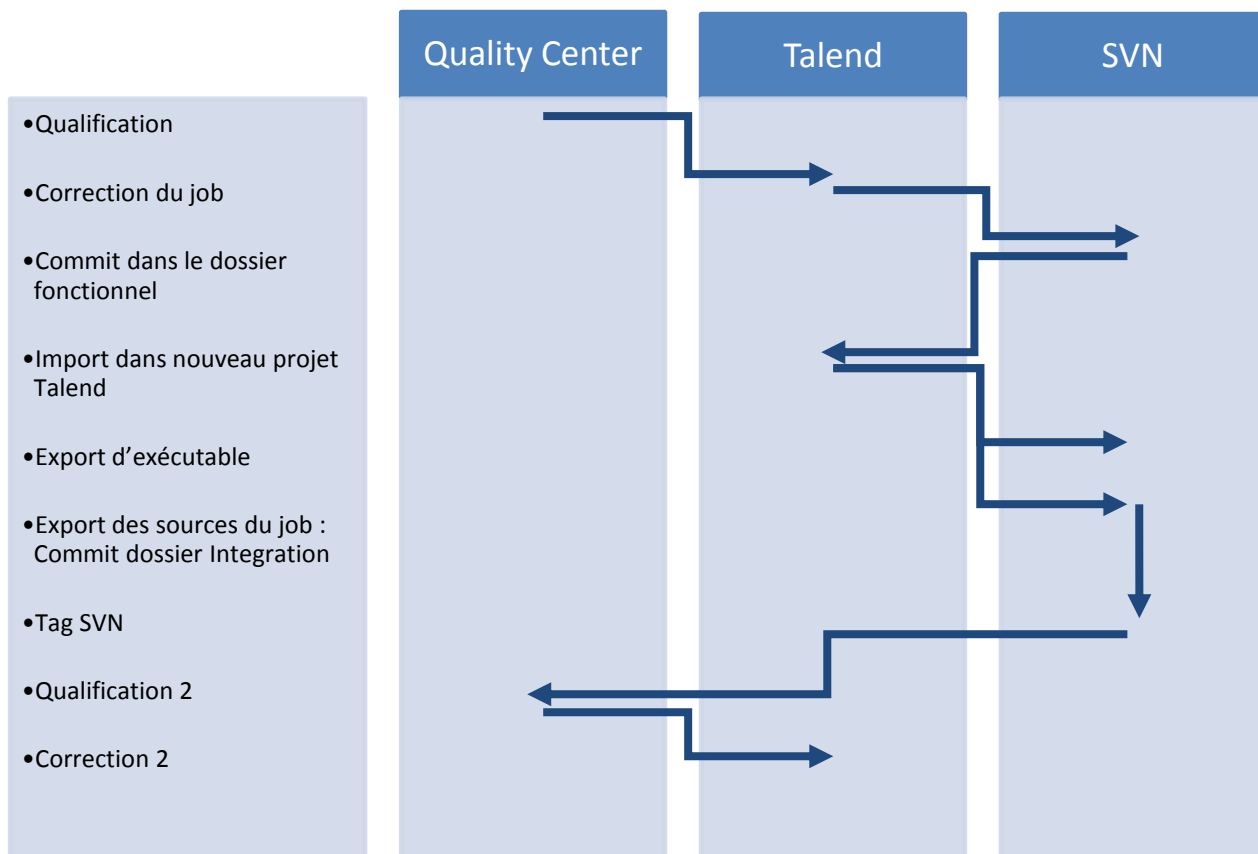
La logique d'une procédure de test au sein de Quality Center est synthétisée dans le schéma suivant :



### 6.4.2. Le principe de Qualification

Dans la méthodologie de travail de notre projet, nous incluons dans la phase de Qualification, à la fois la qualification sur Quality Center mais aussi la correction des anomalies dans l'interface de Talend ainsi que l'intégration du travail individuel dans l'espace partagé avec SVN.

Les différentes étapes de la Qualification et la procédure associée sont résumées par le schéma de processus suivant :



## 7. Conclusion

---

Ce rapport a donc présenté une vue d'ensemble sur mes activités durant mon stage de fin d'études. Ce dernier a été enrichissant sur différents aspects.

D'une part sur les différents sujets techniques abordés et le recours à de multiples technologies. J'ai pu ainsi participer à la mise en œuvre de processus des parties Editique et Reprise de données. Par ailleurs, grâce aux différentes formations, j'ai pu monter en compétence sur plusieurs technologies impliquées dans ce projet.

Ensuite sur le plan organisationnel d'un projet informatique de grande envergure. J'ai pu ainsi découvrir une méthodologie de gestion de projet, à savoir la répartition des tâches, le contrôle et le suivi à un rythme hebdomadaire ainsi que l'interaction entre les différents composants et entre les deux équipes du projet.

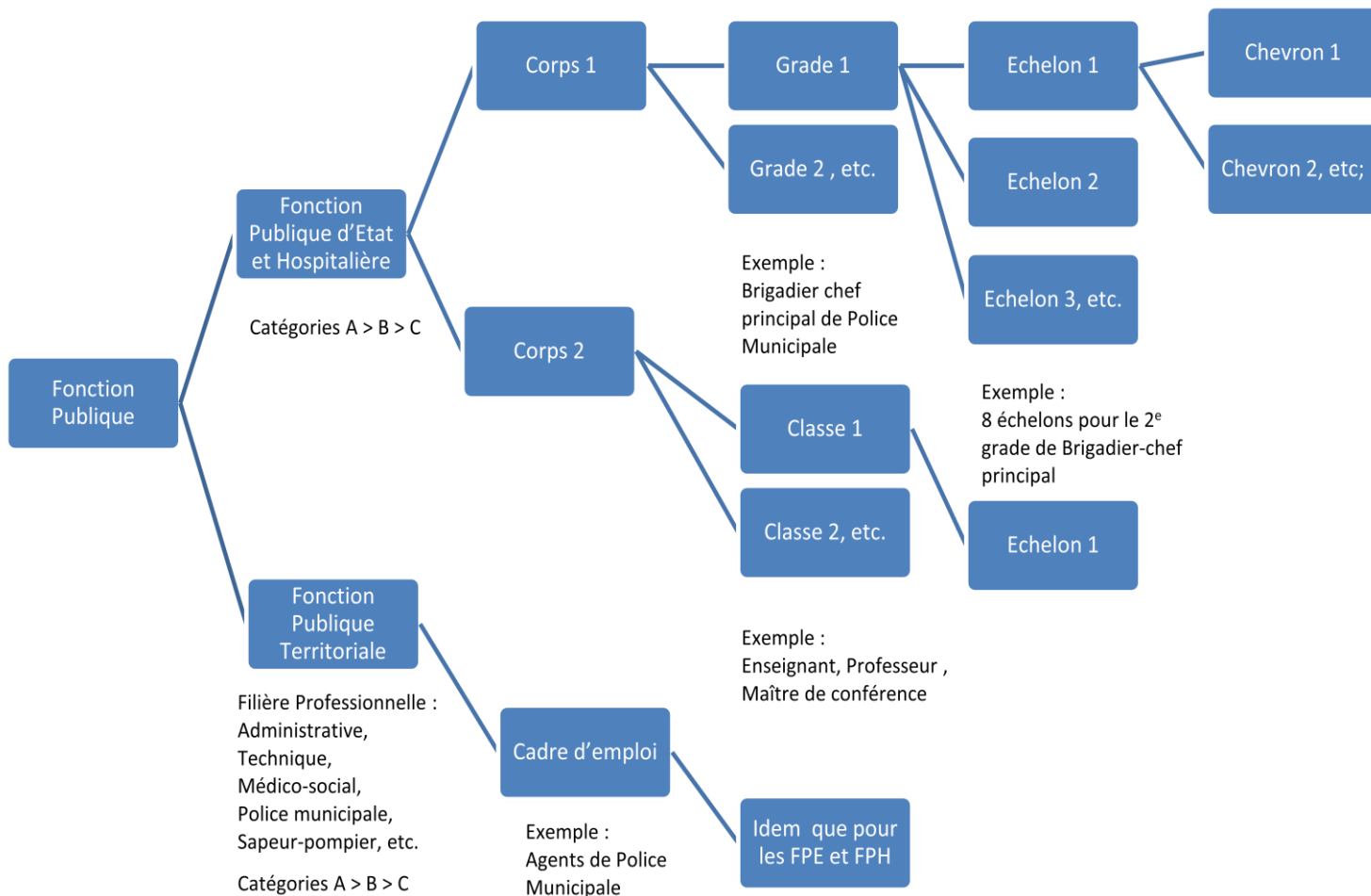
Enfin sur le plan relationnel. Les collaborateurs de Sopra sont généralement jeunes et ont des parcours et des approches uniques. Les différents responsables managériales ou techniques côtoient à un même degré les jeunes collaborateurs. L'échange d'idées et de points de vue avec eux est très formateur.

Par ailleurs, le choix d'effectuer mon stage auprès de Sopra, une ESN, était motivé par une perspective d'embauche et d'évolution professionnelle ainsi que la possibilité d'acquérir de multiples compétences en Informatique et Système d'Information et de monter en compétence sur les aspects techniques, relationnelles voire managériales

Ce stage s'inscrit donc en cohérence avec le parcours de spécialisation en 3<sup>e</sup> Année à l'école Centrale Marseille et plus généralement en cohérence avec la formation professionnalisant de Centrale Marseille.

## 8. Annexes

### 8.1. La Fonction Publique



#### Schéma de l'organisation de la carrière dans la Fonction Publique

La gestion des ressources humaines dans la Fonction Publique repose sur cette organisation de la carrière. Les agents de la Ville de Marseille sont gérés dans le cadre de la Fonction Publique Territoriale.

## 8.2. Les formations reçues dans le cadre du projet

### 8.2.1. La solution éditique APPIC

J'ai pu participer à la première journée de formation, durant laquelle le formateur a fait une introduction à la suite StarJet APPIC avec une présentation des différents modules ainsi que de la logique d'ensemble de l'outil. On a ensuite travaillé à une prise en main du module StarJet Design.

### 8.2.2. Le module Pléiades de gestion de la paie et des temps d'activité PTA

Durant cette journée de formation, j'ai pu découvrir l'organisation du module Paie et Temps d'Activité sur le client lourd et le client léger, ainsi qu'une introduction au langage propriétaire PTA de Pléiades.

### 8.2.3. L'ETL Talend

La formation d'une demi-journée a permis d'introduire l'outil d'EL Talend, de le prendre en main et de manipuler les premiers composants : composants d'affichage, de lecture, d'écriture et de jointure.

### 8.2.4. L'outil de contrôle de tests Quality Center HP

L'outil a été présenté dans les grandes lignes dont principalement le plan de tests, les étapes de test, les vérifications entre les résultats obtenus et ceux attendus ainsi que l'émission d'anomalies (defects).

La prise en main de l'outil s'est faite par la suite au cours de la campagne de qualification.

### 8.2.5. La solution base de données RH de Sopra : Pléiades

Durant cette formation de quatre jours, le client lourd (la station de modélisation) et le client léger (l'application ou portail) ont été présentés ainsi que leur interaction. La modification du modèle de données, à savoir la création de nouvelles tables entités et la création de dépendance entre elles, la création de vues et le déploiement de l'application ont également été abordées à travers des exercices.

Nous avons également pu voir comment manipuler l'outil en tant qu'utilisateur.

Enfin, une introduction au modèle public de données a également été faite.

### **8.3. Autres connaissances mobilisées**

#### 8.3.1. Les macros VBA

J'ai eu l'occasion de me former sur VBA Excel en réalisant des macros pour la partie Éditique.

#### 8.3.2. Le langage Java

Les connaissances acquises à l'école sur le langage Java sont utiles pour la partie Reprise de données.

#### 8.3.3. La conception MCD Modèle Conceptuel de Données

Les connaissances MCD MERISE sont utiles surtout pour la conception des tables de l'éditique.

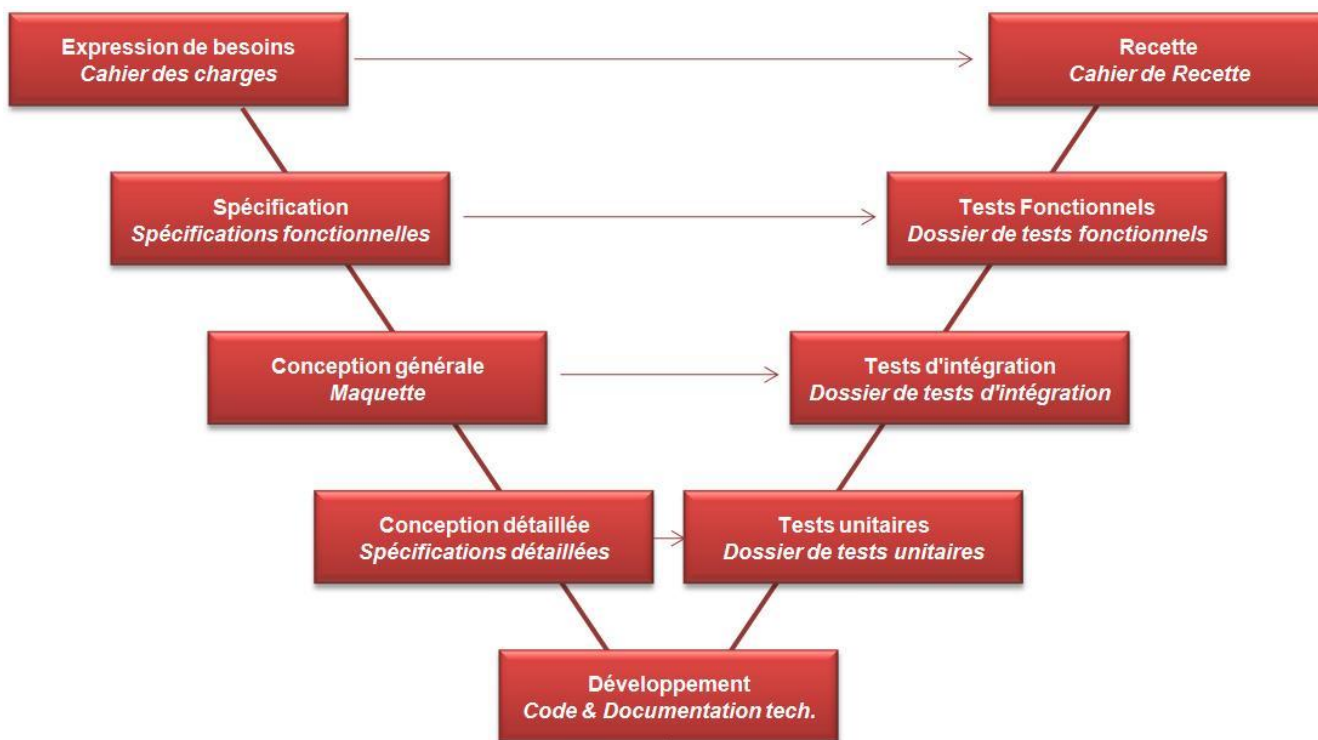
#### 8.3.4. Les connaissances en SQL

Les connaissances en SQL sont surtout utilisées dans la partie Reprise de Données.

Il s'agit en général de jointures entre deux tables, des jointures internes et externes gauches.

On réalise également des sélections SQL.

### 8.4. Cycle de vie en V incluant les tests



Ce schéma d'organisation des phases de développements/tests est globalement suivi sur les deux modules Éditeur et Reprise de données.

En ce qui concerne la partie Éditeur, les quatre premiers points de la branche de droite sont en travail.

Pour la partie Reprise des données, les tests d'intégration sont en cours de finalisation en ce qui concernent les données du premier périmètre.

## 8.5. Principales Références Clients de la Division Sud Est

Banque						
Défense & Aéronautique						
Services, Transport & Utilities						
Industrie						
Protection sociale & Emploi						
Secteur Public						
Énergie						



## 8.6. Bibliographie

### *Documentation et Spécifications*

#### **Documents internes**

Fonctionnel Secteur Public, document PowerPoint.  
Analyse\_docs\_VDM\_Editique\_V2, document PowerPoint.  
SFD Globale - Organisation du processus éditique - V2, document Word.  
SFD Technique APPIC StarJetDesign V2, document Word.  
Processus de recuperation des articles et des visas, document word.  
SFT-Talend-Reprise\_de\_donnees, document Word.  
VDM AZUR SIR Liste des éditions DRH 3, document Excel.

Documents et supports des différentes formations.  
Autres documents internes du projet.

#### **Documents externes**

Fonction publique – fiche n° 5 : le déroulement de carrière, La Gazette.  
L'organisation générale de la fonction publique, emploipublic.fr.

### *Support et Aide sur Internet*

Talend Help Center :

<https://help.talend.com/display/HOME/Welcome>

Stackoverflow :

<http://stackoverflow.com/>

Developpez.com :

<http://www.developpez.com/>

Commentçamarche.net :

<http://www.commentcamarche.net/forum/new>