

Jean-Nicolas HUYGHE  
3ème Année Formation ITI ENSEA  
(Ingénieur en Technique de l'Industrie)  
Présentation PFE  
26 Juin 2020



# Génération automatique d'interfaces homme information pour la Gouvernance et la Supervision d'un Système d'Information dans le domaine de l'Industrie

Contribution au projet MAGISTHER

EDF INGEUM



Tuteur Ecole : Jean-Michel DUMAS

Tuteur Entreprise : Samuel PARFOURU

## Remerciements

Je tiens à remercier mon tuteur Samuel PARFOURU pour son encadrement et pour le savoir qu'il m'a transmis au cours de ces 3 dernières années.

Je remercie également Frédéric GORGETTE et Laura MARANG, mes managers qui m'ont accueilli dans l'équipe d'Informatique Industrielle d'Ingeum.

Je remercie Nicolas DEBROIS, notre chef de projet, pour m'avoir impliqué dans la vie de l'équipe et dans différents projets.

Je suis reconnaissant envers l'équipe 2i pour leur accueil. L'équipe m'a permis d'apprendre auprès de professionnels, de mieux comprendre les enjeux de notre travail et d'améliorer mes compétences techniques mais également relationnelles.

Je remercie également l'équipe e-monitoring et particulièrement Valentin LOUBIERE pour m'avoir accordé du temps afin de mieux comprendre leurs métiers.

Je remercie les prestataires de notre équipe notamment Pascal MAHE, Wajdi BEN-MESSAOUD, Julien N'GOULOU et Nabil NGUERAMAN avec qui j'ai eu l'occasion d'énormément travailler.

Je remercie Jean-Michel DUMAS, mon tuteur pédagogique, pour son accompagnement au cours de ma scolarité à l'ENSEA.

Je suis reconnaissant envers mes relecteurs Aurélie MAGNEZ, Valérie HUYGHE, Robin HESLING et Antoine SIMAO qui m'ont donné un point de vue objectif sur ce rapport.

Enfin, je remercie toutes les personnes avec qui j'ai pu collaborer à long ou court termes au sein d'EDF et à l'ENSEA et qui ont permis d'étoffer mon éventail de connaissances et de compétences.

## Sommaire

|       |  |    |
|-------|--|----|
| 1     | Introduction .....   | 5  |
| 2     | EDF et la transition numérique .....   | 6  |
| 2.1   | EDF SA.....  | 6  |
| 2.1.1 | EDF Opérateur .....  | 7  |
| 2.1.2 | EDF Architecte Ensemblier.....   | 9  |
| 2.2   | Industrie 4.0, Transition Numérique et Activité d'Ingénierie .....                               | 13 |
| 3     | Problématique.....   | 15 |
| 3.1   | L'informatique industrielle et la donnée de process.....   | 15 |
| 3.2   | Les orientations stratégiques d'EDF.....   | 18 |
| 3.3   | Du besoin de Gouvernance de la donnée.....   | 20 |
| 3.4   | Gouvernance, SI et besoin de cartographie .....  | 21 |
| 3.4.1 | Le système d'information et modèle en couches .....  | 21 |
| 3.4.2 | Maintenir une cartographie du système d'information : « le châtimement de Sisyphé » .....        | 21 |
| 3.4.3 | La cartographie du SI comme produit d'un calcul .....  | 22 |
| 4     | La démarche MAGISTHER.....   | 24 |
| 4.1   | CONCEPT .....  | 24 |
| 4.2   | Plateforme MAGISTHER .....   | 27 |
| 5     | MAGISTHER le projet fil rouge de mon alternance .....  | 31 |
| 5.1   | Pilotage de la capture des métadonnées .....   | 31 |
| 5.2   | Le pré traitement des métadonnées via la conception de chaînes ETL.....                          | 32 |
| 5.3   | La cartographie sous la forme d'un graphe de connaissances.....                                  | 33 |
| 6     | De la gouvernance de la donnée à la supervision du SI : extension du périmètre de MAGISTHER..... | 38 |
| 6.1   | Etendre les concepts au sein du graphe de connaissance .....                                     | 39 |
| 6.2   | Générer des interfaces de supervision .....  | 42 |
| 6.2.1 | Contexte .....   | 42 |
| 6.2.2 | Problématique.....   | 43 |
| 6.2.3 | Solution .....   | 43 |
| 6.3   | Indexer la documentation .....   | 48 |
| 6.3.1 | Contexte .....   | 48 |
| 6.3.2 | Problématique.....   | 48 |
| 6.3.3 | Solution .....   | 48 |
| 6.4   | Générer de la documentation.....   | 52 |

|       |   |    |
|-------|---|----|
| 6.4.1 | Contexte .....  | 52 |
| 6.4.2 | Problématique.....                                      | 52 |
| 6.4.3 | Solution .....  | 52 |
| 7     | Discussion & Perspectives .....                         | 54 |
| 7.1   | Bilan.....  | 54 |
| 7.2   | Prochaines Etapes .....                                 | 54 |
| 7.2.1 | Vérification des serveurs automatisé par le graphe..... | 54 |
| 7.2.2 | Supervision des scripts ETL.....                        | 55 |
| 7.2.3 | Analyse d'impact.....                                   | 56 |
| 7.2.4 | Plan du SI .....  | 58 |
| 8     | Conclusion.....   | 60 |
| 8.1   | Bilan professionnel .....                               | 60 |
| 8.2   | Bilan personnel .....                                   | 60 |
| 8.3   | Perspective d'avenir .....                              | 61 |
| 9     | Table des figures .....                                 | 62 |

# 1 INTRODUCTION

---

Nous vivons dans un monde où le numérique est omniprésent. De plus en plus de données sont générées et éparpillées que ce soit sur nos machines personnelles ou dans les serveurs des entreprises.

En effet, l'évolution de plus en plus rapide de nos systèmes d'information (SI) qui s'accompagne d'une explosion du volume de données, désignée sous le terme Big Data, dilue, avec le temps, la compréhension que nous avons de nos SI au sein d'entités telles que les entreprises et conduit à en rendre l'exploitation de plus en plus complexe.

Les données sont acheminées par différents canaux, dupliquées et exploitées pour des utilisations diverses constituant ainsi une chaîne d'approvisionnement des données composée de ramifications complexes. Or, il est difficile d'informer un utilisateur qu'une donnée en amont de celle qu'il utilise n'est plus disponible sans une cartographie permettant ainsi de référencer et visualiser l'ensemble de notre SI.

Dans ce contexte, mon projet de dernière année de formation pour le métier d'ingénieur a pour objectif de concevoir et démontrer la faisabilité de développer une tour de contrôle permettant à la fois de faciliter l'accès mais également la compréhension des informations du SI d'EDF Ingeum maintenu à jour en permanence au service de la gouvernance. Et qui permettrait également de superviser l'architecture du site en continu et de manière automatique.

Ce rapport s'articule autour de cinq parties.

- La première partie décrit l'entreprise EDF, ses enjeux, et détaille l'environnement dans lequel ce projet s'est réalisé.
- La deuxième partie introduit la problématique de gestion des connaissances sur le SI et de l'importance des métadonnées.
- La troisième partie décrit l'avancement du projet MAGISTHER avant son extension.
- La quatrième partie récapitule l'ensemble des développements que j'ai engagé et l'ensemble des solutions que nous avons trouvées afin de pallier aux problématiques évoquées précédemment.
- Enfin, dans une cinquième partie, nous explorerons les possibilités d'améliorations et le futur de ce projet.

Le fil rouge de ce rapport sera donc la gestion de l'information et des connaissances du système d'information d'une entreprise au service de son exploitation.

## 2 EDF ET LA TRANSITION NUMERIQUE

---

### 2.1 EDF SA

EDF (Électricité De France) est une entreprise de production et de distribution d'électricité fondée en 1946 par Marcel Paul. Elle est aujourd'hui, le plus gros producteur d'électricité en France et en Europe et second au niveau mondial (depuis la fusion de China Guodian Group et Shenhua Group).

L'entreprise est actuellement dirigée par Jean Bernard Levy et a réalisé, en 2018, 69 milliards d'euros de chiffre d'affaires. EDF emploie plus de 165 000 personnes.

EDF adresse de grands secteurs d'activité : Opérateur et Architecte Ensemblier.



#### **EDF en tant qu'Opérateur est responsable de la sécurité et du coût de ses usines.**

- Besoin de bien comprendre et de maîtriser techniquement la conception
- Besoin d'intégrer dans la conception le retour d'expérience d'exploitation et les exigences spécifiques au stade de la conception et tout au long de la durée de vie de l'usine.



#### **EDF en tant qu'Architecte Ensemblier n'achète pas de centrales clés en main**

- Spécifie les besoins techniques en fonction des normes et exigence (fonctionnel, performance, sécurité...)
- S'appuie sur des concepteurs responsables (NSSS, CI, I&C,...)
- Examine les études détaillées et contrôle les produits, composants, travaux livrés par la R&D
- Responsable de la démonstration de sécurité de l'usine.

*Figure 1 Les deux grands métiers d'EDF*

## 2.1.1 EDF Opérateur

Un opérateur est une entreprise ayant pour mission d'entretenir et d'exploiter une infrastructure liée à un service public.

EDF opère sur l'ensemble de la chaîne de valeur de l'énergie tel que présenté sur la figure suivante : Production, Négociation, commercialisation et services énergétiques. Elle intervient également dans le transport, la distribution et l'approvisionnement en gaz naturel.

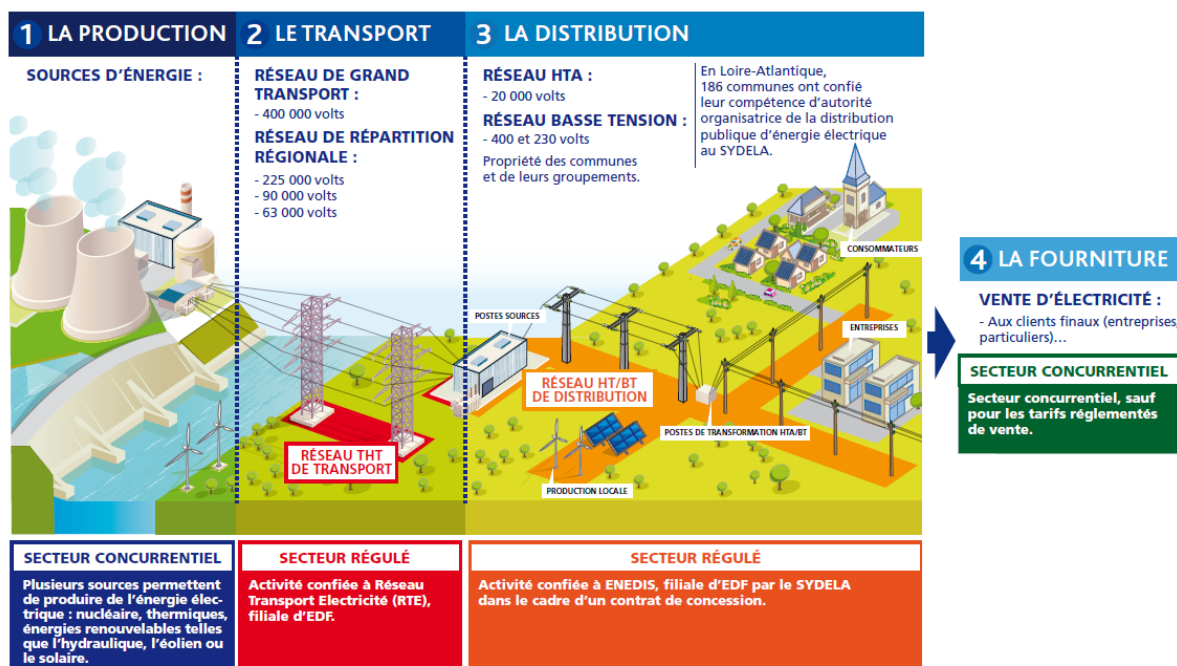


Figure 2 Chaîne de valeur du secteur de l'énergie

Dans le domaine de la production, le groupe EDF a construit et exploite le plus grand parc de production énergétique Européen avec 139 Giga Watt (GB) de capacité installée. Le groupe est très impliqué dans la lutte contre le réchauffement climatique avec une production bas carbone représentant 87% de la production d'électricité du groupe. Cette production permet de répondre à la demande de plus de 39 millions de clients à travers le monde.

### 2.1.1.1 Mix énergétique

Pour qu'un réseau électrique puisse fonctionner et fournir une électricité de bonne qualité, en particulier sur le plan de la stabilité de la fréquence, il est nécessaire d'équilibrer en permanence la production et la consommation.

Le besoin en électricité par contre en France et dans le monde n'est pas constant et seulement partiellement prévisible. Le besoin fluctue selon l'heure de la journée (plus de besoin le soir), selon la météo (plus de besoin quand il fait sombre) mais aussi des jours de la semaine (utilisateur plus chez lui lorsque c'est le week-end ...). La courbe de charge sur la figure suivante illustre ce phénomène sur une journée.

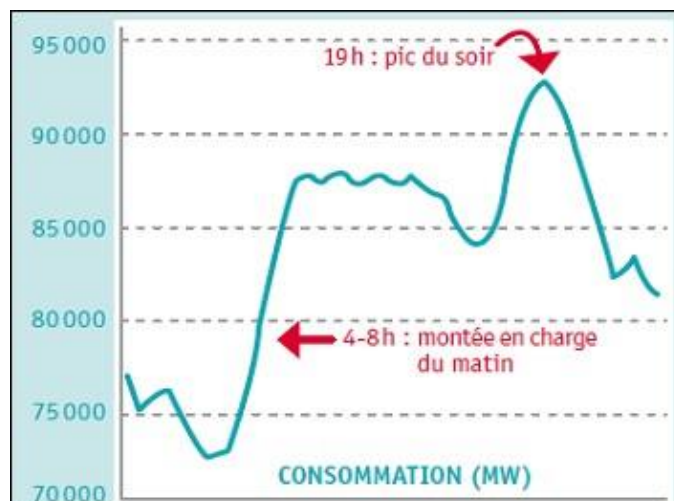


Figure 3 Courbe représentative de la consommation d'électricité moyenne en France pour une journée

Ainsi, pour répondre à l'équilibre entre production et consommation, beaucoup de paramètres entrent en compte pour calculer le besoin en électricité jour après jour et en conséquence adapter la mobilisation des moyens de production sur une période donnée.

Le choix des moyens de production à mobiliser se fait en considérant leur disponibilité mais également leur rapidité plus ou moins importante à alimenter le réseau. Ce explique le recours à un mix énergétique et non pas une filière unique de production d'énergie. La figure suivante illustre la répartition du mix énergétique français.

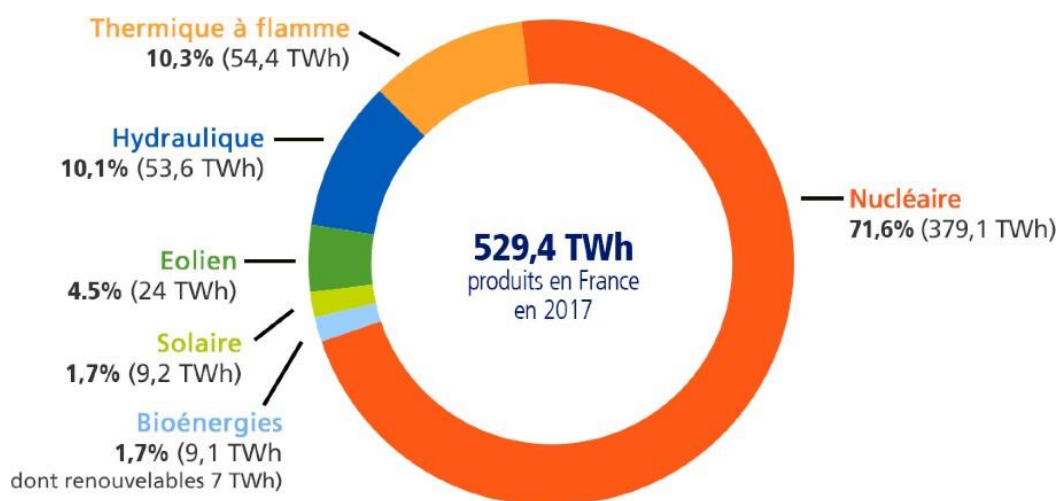


Figure 4 Production Française d'électricité en 2017

Chaque moyen de production est plus ou moins adapté à une situation.

Par exemple, les centrales nucléaires produisent une quantité d'énergie importante et constante mais bien qu'elles soient capables de faire du suivi de charge, elles ne peuvent répondre à des fluctuations trop importantes. Les moyens hydrauliques, en particulier les barrages sont le moyen de production le plus flexible car on peut aisément asservir leurs productions ainsi que stopper ou démarrer leurs fonctionnements quasiment à la demande. La filière thermique en fonction de la technologie employée peut être mobilisée entre 45min, c'est le cas des turbines à combustion, à quelques heures, en fonction de la température de la chaudière par exemple pour des moyens de production de type centrale à charbon ou gaz.



## 2.1.2 EDF Architecte Ensemblier

Par définition, un architecte ensemblier est une entreprise fournissant un service de conception et de construction architectural. Elle intervient également dans la déconstruction de site inactif en développant des techniques innovantes afin de réduire les coûts et le temps de démantèlement des sites.

EDF a une longue tradition d'architecte ensemblier qui, dans son histoire se traduit par exemple par le grand programme hydraulique suite à la Seconde Guerre mondiale, le développement d'un réseau de transport d'énergie sur l'ensemble du territoire français mais également dans le programme nucléaire qui a permis de construire 58 réacteurs nucléaires en moins de 20 ans. Ces savoirs faire sont au cœur des ingénieries (nucléaire, hydraulique, thermique à flamme) d'EDF et sont mobilisés pour le groupe mais également pour des clients en France et à l'international.

EDF fait ainsi valoir son expérience dans la conception et construction de centrales nucléaires avec les chantiers de Flamanville et de Hinckley Point C (Royaume Uni). Ces deux chantiers visent à construire des réacteurs de nouvelle génération permettant une production d'énergie plus performante, plus propre et en conformité avec les plus hauts niveaux de sûreté mondiaux.

Sur le volet actif de production, EDF en tant qu'Architecte Ensemblier est engagé dans les missions suivantes :

- **La conception de nouveau modèle de réacteur** (prenant en compte les faisabilités techniques, les restrictions écologiques et les différents besoins en énergies ...);
- **Entreprendre des chantiers** visant à déployer de nouveau modèle de réacteur ;
- **Le support sur des sites existants** (remplacement de pièces dans les réacteurs, études de durabilité des équipements...)
- Enfin, **le désassemblage de centrales inactives**. En effet, EDF met en place des solutions et stratégies afin de déconstruire les centrales le plus rapidement et efficacement possible.

### 2.1.2.1 La DTEAM & EDF Ingeum

Le maintien du parc thermique à flamme incombe à la DTEAM (Division Thermique Expertise Appui industriel Multi métier). Cette entité réunit à la fois la gestion des moyens de production thermique à flamme existant mais également l'ingénierie thermique. Cette entité inclut également l'ingénierie des réseaux de distribution.

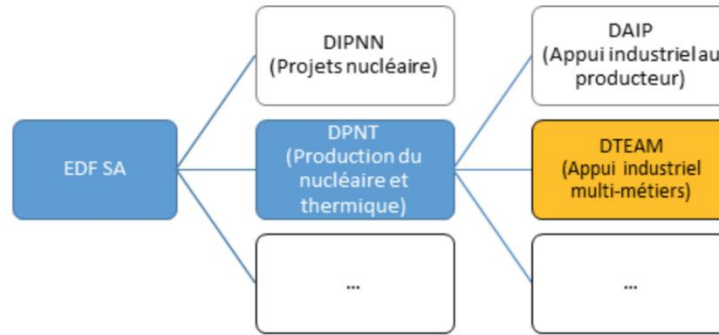


Figure 5 Organigramme de la DTEAM

Le parc thermique à flamme reste une composante nécessaire et indispensable du mix énergétique et fait l'objet d'une constante amélioration technique, économique, tout en s'adaptant à une réglementation environnementale en évolution constante.

Au sein de la DTEAM, l'ingénierie thermique est regroupée au sein d'une entité nommée. Ingeum propose un service d'assistances liées au maintien et à l'exploitation des centrales thermiques pour l'ensemble de ses clients à travers le monde.

Ingeum compte 536 salariés et est découpé en trois lignes :

- La Ligne de projets et service pour les centrales thermiques d'EDF en France ;
- La Ligne de projets et service à l'international ;
- La Ligne de projets et service pour le Groupe EDF : des prestations d'ingénieries pour le nucléaire, Dalkia, EDF EN (Energies Nouvelles), PEI – SEI (Production Electrique Insulaire et Système Electrique Insulaire) production pour les îles françaises (Corse, Antilles, ...).

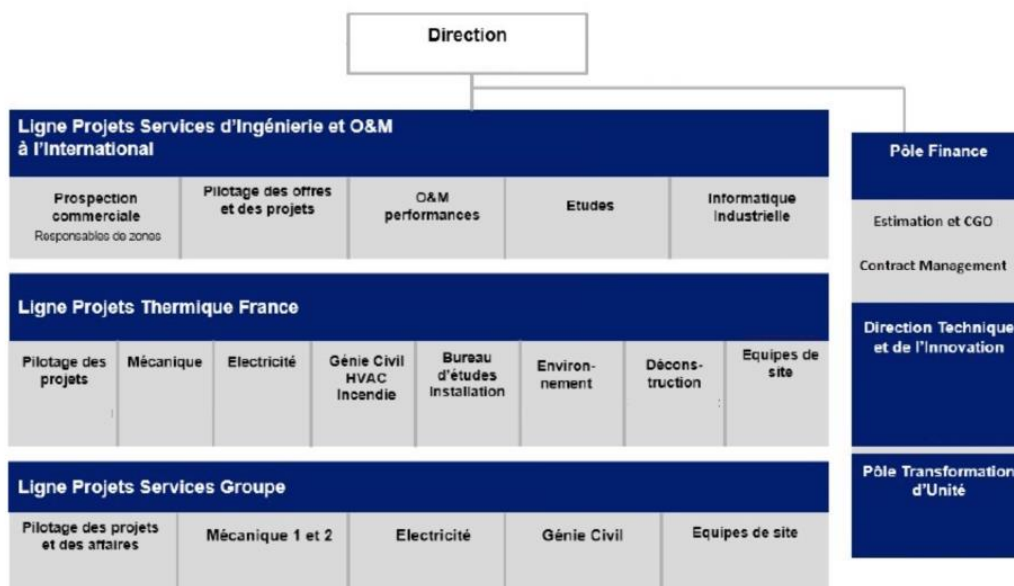


Figure 6 Organigramme EDF Ingeum 2018

Les activités conduites par ses trois lignes sont :

- Les études d'ingénierie et le suivi des travaux de maintenance des centrales thermiques d'EDF SA. Ces projets renvoient à la rénovation des centrales au gaz et au charbon ainsi qu'à la construction des centrales CCG (cycle combiné gaz) en France et dans le monde;
- La déconstruction de centrales en inactivité;
- Le conseil et le service à l'international sur les centrales thermiques à flamme.

### 2.1.2.2 La ligne Internationale

La ligne internationale a une activité de services centrée autour de 3 sujets :

- La construction,
- La production,
- La maintenance des centrales thermiques.

Ces services sont commercialisés dans les zones Moyen Orient, en Afrique, en Europe, en Asie.

Ces activités principales sont basées sur la vente de services à travers 6 axes différents et dont le service d'e-Monitoring qui s'intègre dans l'axe d'exploitation et maintenance.

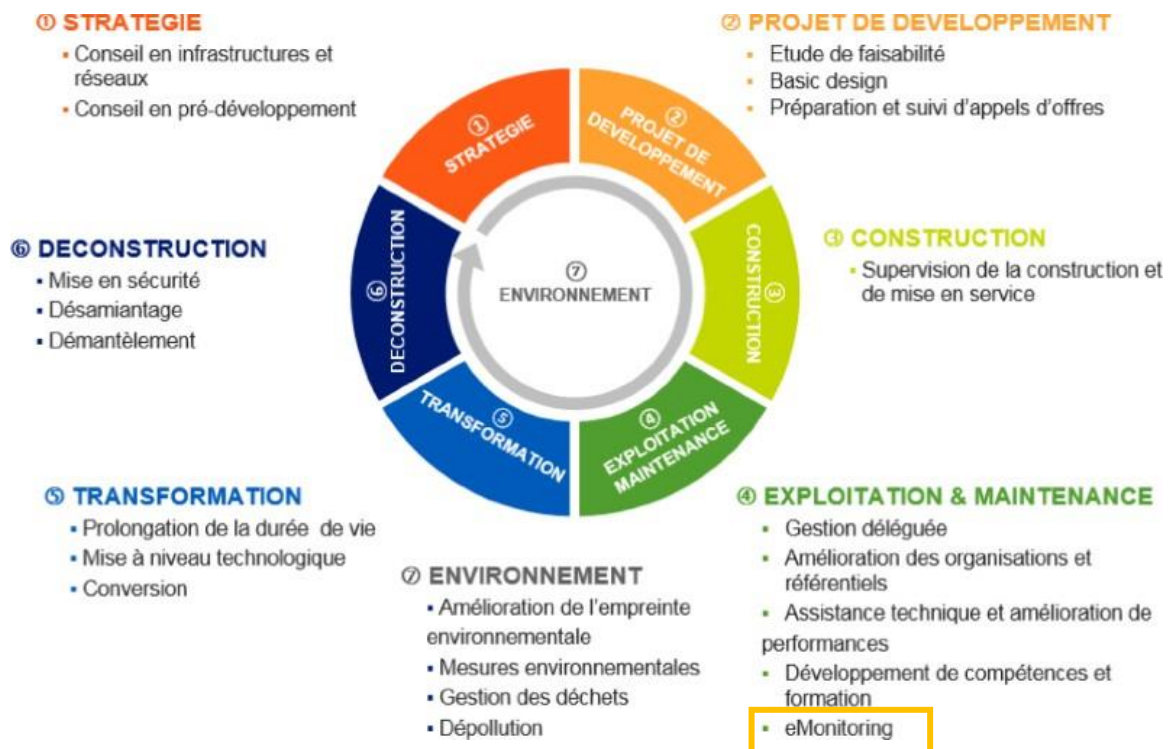


Figure 7 Stratégie EDF Ingeum Ligne Internationale

L'e-Monitoring est une équipe d'une vingtaine d'ingénieurs chargés d'exploiter les données remontées depuis les centrales thermiques afin de fournir des rapports sur l'état de la centrale aux exploitants. Ils sont en charge d'une offre de Détection Précoce d'Anomalie (DPA) dont l'objectif de prévenir les exploitants en amont d'une éventuelle défaillance sur la centrale afin par exemple d'anticiper un changement de pièce.

Les données exploiter par l'e-Monitoring leurs sont rendu accessible par l'équipe Informatique Industrielle. La compétence de l'équipe 2i est indispensable au travail réalisé par l'e-monitoring et pour les services qu'elle fournit.

### 2.1.2.3 *L'équipe Informatique Industrielle*

Au sein de la ligne internationale, j'ai intégré le service d'Informatique Industrielle en 2017.

Ce groupe est composé d'une vingtaine d'ingénieurs et prend en charge les activités suivantes :

- **L'activité informatique industrielle** : sa mission peut se résumer à l'acheminement des données industrielles à partir du moyen de production pour la mettre à disposition de tous les membres de l'ingénierie ou du producteur,
- **L'activité contrôle commandes** : consiste à réaliser l'ingénierie de construction et de maintenance des contrôles de commande de centrale ;
- **L'activité Cyber Sécurité** : consiste à réaliser l'ingénierie de sécurité des systèmes d'informations des centrales ;
- **L'informatique de gestion** : consiste à gérer les moyens informatiques matériels, logiciels et télécom pour le personnel d'EDF Ingeum.
- **La gestion de connaissances** : appuie EDF Ingeum dans la gestion documentaire des projets d'EDF Ingeum dont l'enregistrement en GED (Gestion Electronique des Documents) et l'archivage.

Je contribue et intervins dans l'activité de mise à disposition des données de process de centrales dont un client interne privilégié est l'équipe e-Monitoring.

## 2.2 INDUSTRIE 4.0, TRANSITION NUMERIQUE ET ACTIVITE D'INGENIERIE

« Le concept d'industrie 4.0 ou industrie du futur correspond à une nouvelle façon d'organiser les moyens de production. Cette nouvelle industrie s'affirme comme la convergence du monde virtuel, de la conception numérique, de la gestion (opérations, finance et marketing) avec les produits et objets du monde réel »<sup>1</sup>.

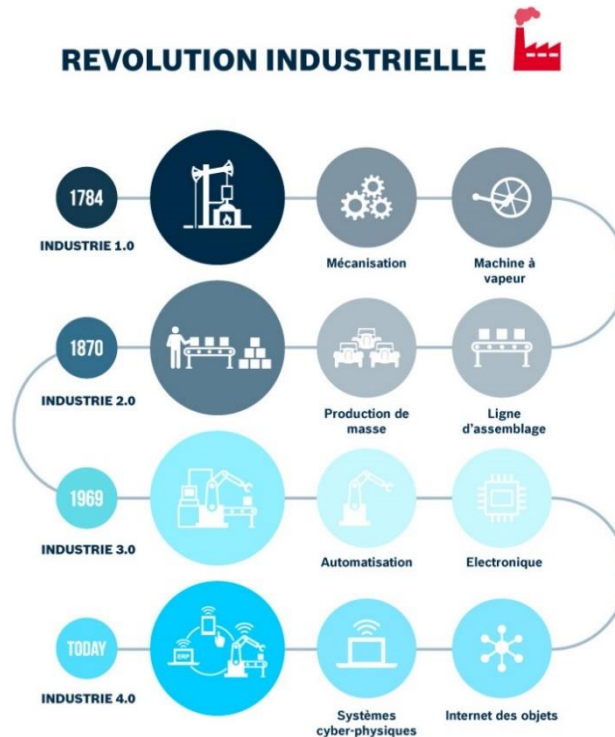


Figure 8 Industrie 4.0 : la 4<sup>ème</sup> révolution Industrielle<sup>2</sup>

Cette 4<sup>ème</sup> révolution industrielle « organise des processus de production induits par les innovations liées à l'internet des objets et aux technologies du numérique, tels que la robotique, la réalité augmentée, l'impression 3D, l'intelligence artificielle, afin d'exploiter les données issues du Big data et de la maquette numérique. Toutes ces techniques font partie des ingrédients qui entrent dans la composition de l'Industrie 4.0 »<sup>3</sup>.

La transition numérique est au cœur de cette vision Industrie 4.0. Elle se fonde sur les principes suivants :

- Travailler autrement avec un ensemble d'outils collaboratifs ouvert sur le monde qui vise à rapprocher l'humain et l'amener dans des démarches co-créatives et productives,
- Décloisonner les organisations et le management traditionnel au profit d'un management transversal.

<sup>1</sup>

Source : [https://fr.wikipedia.org/wiki/Industrie\\_4.0#:~:text=Le%20concept%20d'industrie%204.0,et%20objets%20du%20monde%20r%C3%A9el.](https://fr.wikipedia.org/wiki/Industrie_4.0#:~:text=Le%20concept%20d'industrie%204.0,et%20objets%20du%20monde%20r%C3%A9el.)

<sup>2</sup> Source : <https://expertise.boschrexroth.fr/tout-comprendre-de-industrie-4-0/>

<sup>3</sup>

Source : [https://fr.wikipedia.org/wiki/Industrie\\_4.0#:~:text=Le%20concept%20d'industrie%204.0,et%20objets%20du%20monde%20r%C3%A9el.](https://fr.wikipedia.org/wiki/Industrie_4.0#:~:text=Le%20concept%20d'industrie%204.0,et%20objets%20du%20monde%20r%C3%A9el.)

Les technologies matériels et logiciels sont la clé de voûte pour produire le flux, le carburant, nécessaire à cette transition : la data. La data est un levier, un support, permettant d'organiser le travail autrement.

Dans le secteur de l'énergie et l'ingénierie en particulier la donnée est partout sous des modalités hétérogènes. Ainsi, l'enjeu pour une ingénierie est sur l'ensemble du cycle de vie d'un actif de production, de contribuer à toutes phases, gérer et maintenir les compétences et connaissances et pour cela il est indispensable de gérer l'information et la connaissance.



Figure 9 L'ingénierie au cœur de la gestion de l'information sur l'ensemble du cycle de vie d'un actif de production

Dans toutes ces données à maintenir et à exploiter, il y a également les données de process dont la responsabilité incombe au service d'informatique Industrielle.

### 3 PROBLEMATIQUE

#### 3.1 L'INFORMATIQUE INDUSTRIELLE ET LA DONNEE DE PROCESS

Le groupe Informatique Industrielle d'EDF INGEUM est un acteur clé de la performance des centrales thermiques et de la compétitivité des prestations de conseil d'EDF INGEUM en France et à l'international.

De par ses activités, l'équipe Informatique industrielle d'INGEUM contribue à la conservation et la valorisation de données. Elle exploite et continue à développer une infrastructure de gestion en flux des données du process en assurant :

- L'acquisition (en particulier via les interfaces avec le Système Numérique de contrôle commande des centrales de production (environ 200 000 variables acquises toutes les secondes sur une trentaine de sites répartis sur plusieurs continents),
- Le transport (architecture réseau prenant en compte les risques cyber sécurité d'intrusion sur les SNCC),
- Le stockage et traitement,
- La mise à disposition (via des applications expertes (logiciels sur étagère), un portail utilisateur (développement interne), une plateforme d'analyse avec des solutions de type Business Intelligence (BI), Data Lake, des API, etc...)
- La gouvernance (urbanisation) de l'ensemble de l'infrastructure (données de process et métadonnées)

Le profil typique de chaîne d'acquisition de données actuellement en place est schématisé sur la figure suivante (il s'agit d'une version simplifiée qui ne traduit pas tous les détails).

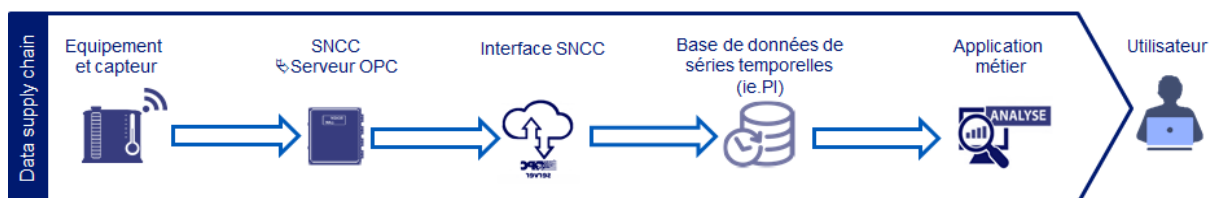


Figure 10 La chaîne d'approvisionnement des données

Les données de capteurs sont acheminées et transitent via une succession d'organes qui constituent alors une chaîne, un réseau, d'acheminement de données. Le point de départ de la chaîne est le capteur dont le signal va remonter, via des interfaces ou automates, vers un contrôle commande numérique, le SNCC pour Système Numérique de Contrôle Commande. Les SNCCs modernes offrent des interfaces permettant d'exposer leurs données et de les pousser vers des systèmes en aval en exploitant en particulier la famille de protocole OPC<sup>4</sup>. Dans ce type d'intégration, la donnée est poussée vers des interfaces en aval qui sont « à l'écoute ». Pour l'infrastructure d'INGEUM, la donnée transite ensuite vers des bases de données de séries temporelles (i.e. Historian<sup>5</sup>) dans le cas en présence un PI System<sup>6</sup> de la société Osisoft.

<sup>4</sup> Source : [https://en.wikipedia.org/wiki/Open\\_Platform\\_Communications](https://en.wikipedia.org/wiki/Open_Platform_Communications)

<sup>5</sup> Source : [https://en.wikipedia.org/wiki/Operational\\_historian](https://en.wikipedia.org/wiki/Operational_historian)

<sup>66</sup> Source : <https://www.osisoft.com/pi-system/>

L'Historian constitue une base de données patrimoniale qui permet de rendre disponibles les données auprès d'ingénieurs via une application ou des dispositifs artificiels (algorithme, calculateur...) afin d'être exploitées et valorisées.

Les différents maillons de la chaîne sont à la fois une nécessité mais également une source potentielle de problèmes pouvant compromettre l'acheminement et la valorisation de la donnée. La figure suivante présente quelques situations pouvant apparaître à différents niveaux du cheminement.

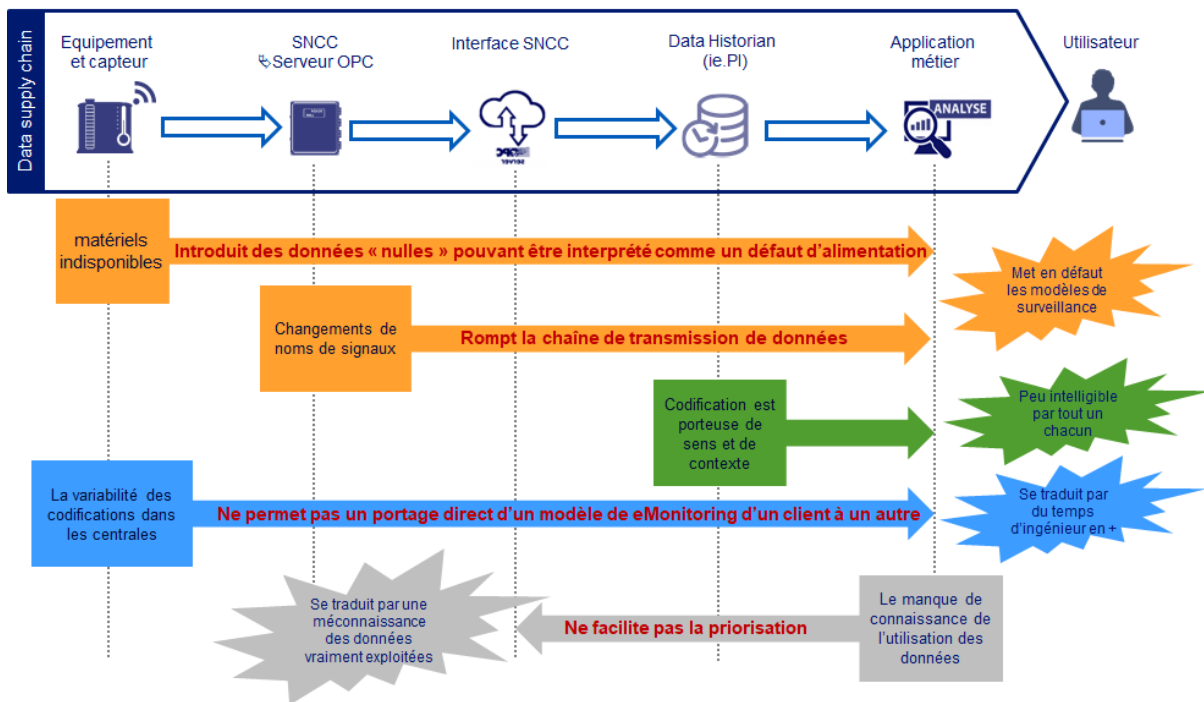


Figure 11 Les problèmes liés à la chaîne d'approvisionnement des données

Pour que la chaîne fonctionne correctement un élément fondamental est la notion de TAG qui référence les séries temporelles de manière unique. Chaque interface de transit ou base de données de stockage possède son plan de référencement. D'un équipement, à l'autre ses plans peuvent être identiques mais parfois on introduit des transformations sous la forme de préfixe, suffixe ou autres subtilités.



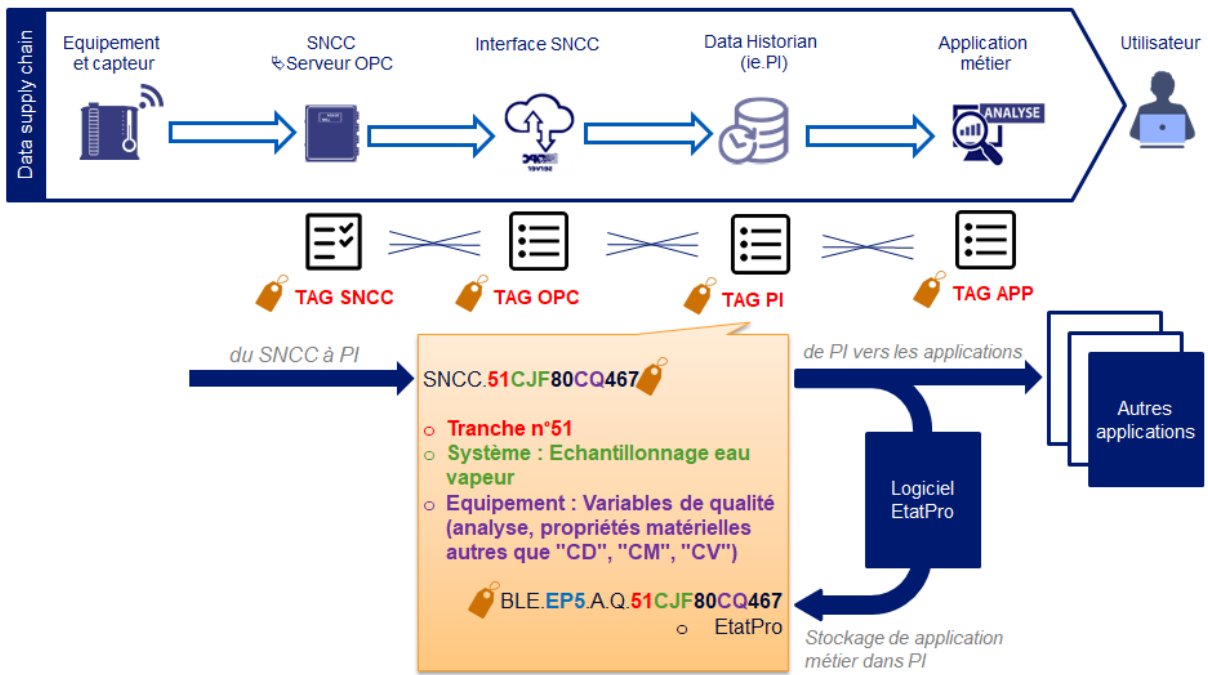


Figure 12 Notion de TAG

Aussi, parmi les problématiques majeures, on peut citer de lignage des TAGs entre les différents organes d'acheminement. La donnée est le plus souvent poussée et pour que cela fonctionne correctement un mapping doit être maintenant en permanence sur l'ensemble de la chaîne.



Figure 13 Mapping de tag

Dans l'exemple ci-dessous, on voit un TAG en surbrillance (en rouge sur l'illustration) qui est utilisé dans une application métier. Il est alimenté par un TAG PI de la stack Monither (Orange sur l'illustration). Ce dernier est alimenté par un TAG PI de la stack de Martigues (orange sur l'illustration dans la stack de Martigues). Enfin, il existe en amont de tous ces points, un TAG OPC (en jaune sur l'illustration) qui alimente le TAG suivant. A chaque étape de la chaîne d'approvisionnement, les TAGs gardent un radical commun préfixé en fonction de cette dernière. Aujourd'hui, il n'existe pas de solution pour réaliser le mapping de toutes ces variables afin de retrouver la chaîne d'approvisionnement complète.

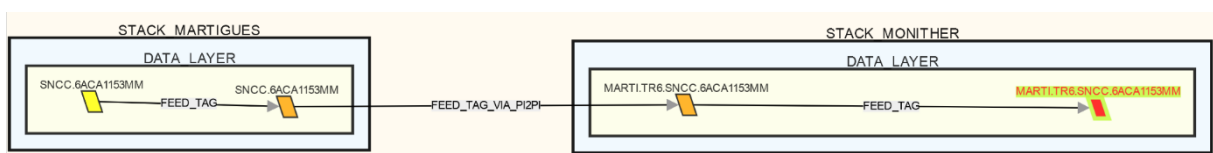


Figure 14 Exemple de reconstitution de la chaîne d'acheminement des données

Si par exemple, le référencement d'une série temporelle change de TAG à la source, l'aval de la chaîne ne sera pas alimenté si le mapping n'est pas réadapté. Il est donc crucial de capitaliser et modéliser l'ensemble de la chaîne d'acheminement de données et en particulier les mappings

entre TAGs d'équipements voisins afin de pouvoir superviser le système global, le maintenir et adapter les configurations face à des changements subits et non anticipés ou planifiés.

Pour ce faire, il faut Inventorier, Cataloguer, et Modéliser l'infrastructure de données selon différents points vus et niveaux afin de pouvoir gouverner le système dans son ensemble.

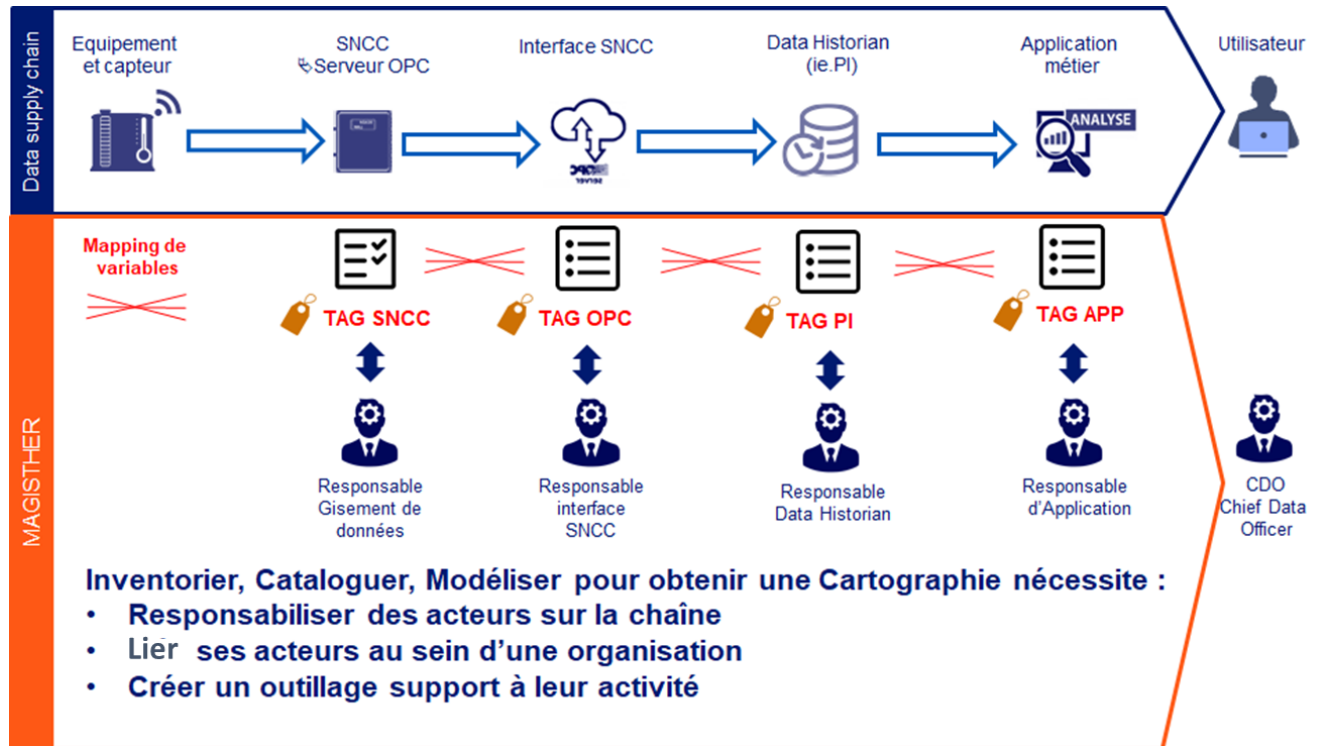


Figure 15 Magisther et la chaîne d'approvisionnement

Cela doit conduire à responsabiliser des acteurs sur l'ensemble de la chaîne, les lier au sein d'une organisation et créer un outillage support à leur activité.

### 3.2 LES ORIENTATIONS STRATEGIQUES D'EDF

Le plan Cap 2030 présente les orientations stratégiques d'EDF. La figure synthétise le plan Cap 2030 :

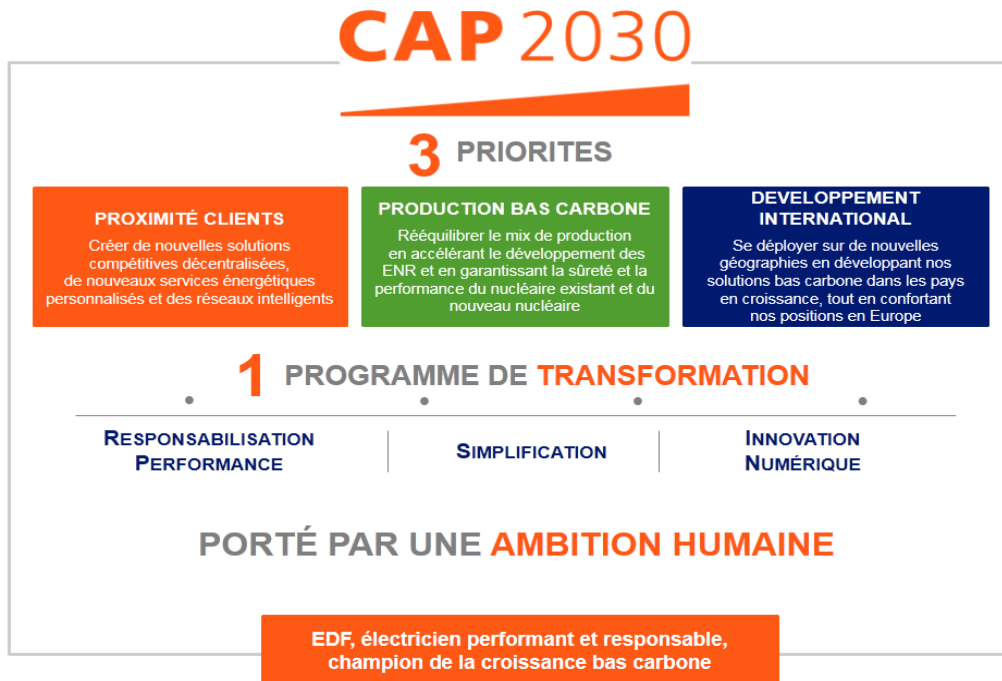


Figure 16 : Image provenant du document officiel de la présentation du cap 2030

Ce plan définit trois priorités :

- **La Proximité avec le client** : être plus compétitif, créer un réseau plus intelligent ;
- **La Production du bas carbone** : développer les énergies renouvelables, réduire les gaz à effet de serre des centrales thermiques, le grand carénage visant à allonger de la durée de vies des centrales nucléaires en intégrant les améliorations du niveau de sûreté (tenant compte du retour d'expérience de Fukushima) ;
- **Le Développement à l'international** : gagner de plus en plus de projets de création de centrales ou des projets d'énergies renouvelables à l'étranger.

Pour répondre à ces 3 priorités, EDF s'est engagée dans un vaste projet de transformation visant à :

- Simplifier, innover, responsabiliser et faire évoluer les pratiques managériales ;
- Digitaliser les processus traditionnels pour les rendre plus performants.

On voit ici que dans cette transformation le numérique et le système d'information doit jouer un rôle de premier plan car ils sont au cœur de la circulation de la donnée qui doit être gérée « comme un **actif stratégique** au service de la **valorisation** » telle qu'évoquée par le **comité e-Transformation EDF Groupe**.

Cette transition vers le numérique accompagne la 4<sup>ème</sup> révolution industrielle amenant aux industries 4.0.

### 3.3 DU BESOIN DE GOUVERNANCE DE LA DONNEE

La gouvernance des données a pour objectif d'assurer l'utilisation efficace et efficiente des données d'une entreprise afin qu'elle mène à bien ses objectifs.

La gouvernance doit permettre :

- Une meilleure compréhension de l'information en adoptant une terminologie commune, accessible à tous (tout en conservant les données techniques pour le métier) ;
- Une qualité des données certifiées par la structuration d'un plan cohérent et exhaustif des données.

Gouverner c'est prévoir et il n'est pas de bonne gouvernance sans avoir connaissance de son environnement. Il est possible de faire une analogie entre la gouvernance et l'utilisation d'un logiciel d'aide à la navigation comme Waze<sup>7</sup>.



Figure 17 Le logiciel Waze

A l'image de ce logiciel qui offre la possibilité de visualiser une carte géographique et indique le chemin à suivre en fonction d'une destination, la gouvernance doit s'appuyer sur une **Cartographie** complète et à jour des données de l'entreprise afin de pouvoir suivre l'acheminement de la donnée ainsi que ses tenants et ses aboutissants.

En l'absence d'une cartographie de ce type, pour un gestionnaire d'infrastructure il est difficile de vraiment être maître de celle-ci et l'on tend à subir le système que l'on doit gérer. En effet, si l'on ne sait pas qu'un flux de données est inopérant comment prévenir un utilisateur en bout de chaîne que son système ou application ne sera pas alimenté sans cartographie de l'ensemble ?

<sup>7</sup> Source : <https://www.waze.com/>

### 3.4 GOUVERNANCE, SI ET BESOIN DE CARTOGRAPHIE

#### 3.4.1 Le système d'information et modèle en couches

Le système d'information est classiquement représenté par un empilement de couches successives qui traduisent différentes visions et différents niveaux d'abstraction. La figure suivante illustre l'ensemble des couches généralement rencontrées auxquelles sont associés sur la gauche les objets métiers principaux et sur la droite les types de référentiels et cartographies généralement fabriqués.

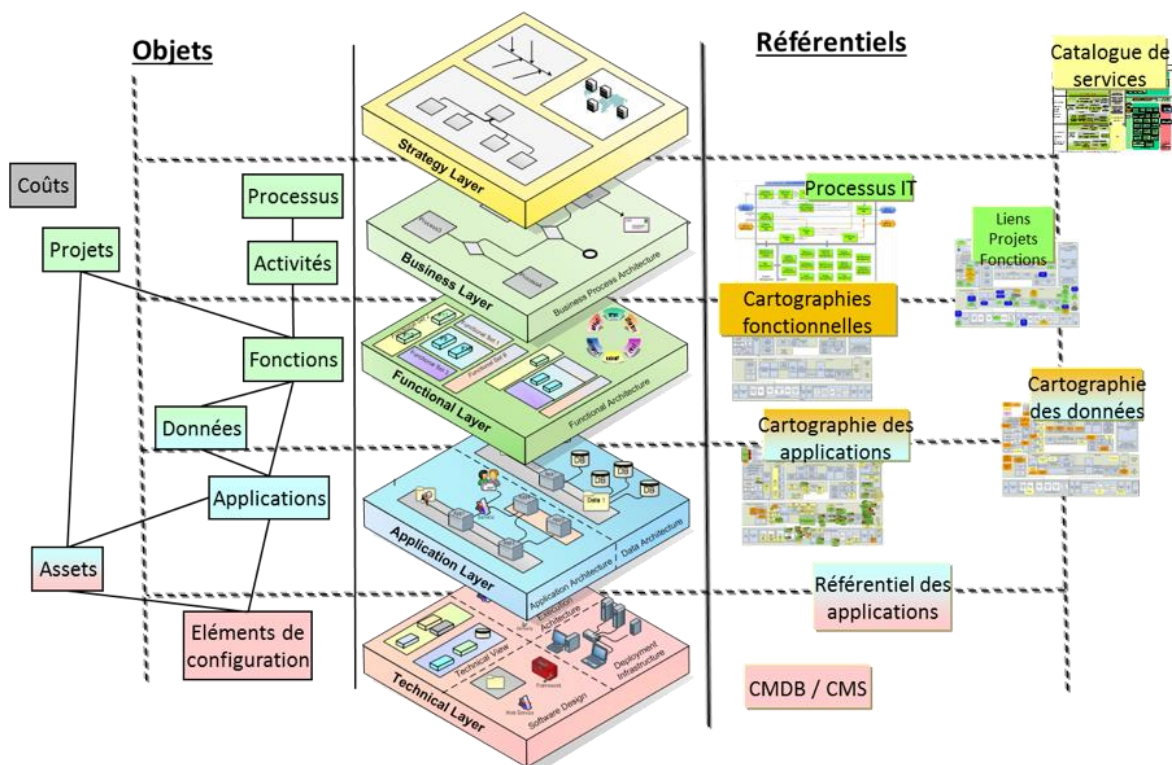


Figure 18 Urbanisation du SI<sup>8</sup>

#### 3.4.2 Maintenir une cartographie du système d'information : « le châtiement de Sisyphe »

Beaucoup d'acteurs de spécialités différentes concourent à la fabrication et l'alimentation des référentiels du système d'information. Evoquons de manière non exhaustive les architectes d'entreprise et les urbanistes et plus généralement les gestionnaires de l'ensemble des objets qui font le système d'information. Le maintien de ces référentiels est une activité au long cours qui demande des efforts constants pour inventorier et modéliser les relations entre les différents objets du système d'information.

La fragmentation et l'accélération de l'évolution de nos systèmes d'informations alimentées par des phénomènes comme le Big Data, l'IoT, l'approche Data-Centric et plus généralement

<sup>8</sup> Source : <https://www.supinfo.com/articles/single/2987-urbanisation-systemes-information>

le mouvement de fond de transition numérique tend à endiguer la faisabilité réelle de cartographier le système d'information.



Figure 19 Le rocher de Sisyphe<sup>9</sup> : de la mythologie à une vision à l'ère moderne

En effet, dans ce contexte, maintenir un référentiel et une cartographie du système d'information peut aisément s'apparenter au « châtiement de Sisyphe<sup>10</sup> » dans le sens où à peine le référentiel est-il rédigé, modélisé, il se révèle déjà obsolète éventuellement erroné.

Cette difficulté est bien évidemment plus ou moins notable en fonction de la couche du SI que l'on cherche à représenter. Certaines parties ont une cinétique d'évolution lente. Cependant à l'heure d'une virtualisation de plus en plus importante du SI avec en particulier la montée en puissance de la virtualisation applicative et des architectures micro-services via la démocratisation croissante de la technologie Docker, le besoin de plus en plus important de tracer et justifier la prolifération et les usages faits de toutes données que cela soit pour répondre à la menace cyber ou la mise en conformité à des règlements tel que le RGPD, il semble inéluctable de systématiser et automatiser le processus de cartographie du système d'information.

### 3.4.3 La cartographie du SI comme produit d'un calcul

Chaque couche ou point de vue sur le SI se traduit en général par des référentiels indépendants qui ne référencent pas les mêmes types d'objets. La figure suivante illustre la diversité des objets métiers. Aussi, pour outiller une gouvernance, l'enjeu n'est pas de se limiter à créer des référentiels et des cartographies par couches mais bien de mailler des référentiels multiples entre eux afin d'envisager que les cartographies ne soient que le produit d'un calcul basé sur ses référentiels maillés.

<sup>9</sup> Source : <http://remacle.org/bloodwolf/liege1/Fantasmagories/1fan2.htm>

<sup>10</sup> « Sisyphe est surtout connu pour son châtiement, consistant à pousser une pierre au sommet d'une montagne, d'où elle finit toujours par retomber. » [Wikipédia]

- Vision
  - Objectifs, Risques, Exigences, Glossaire,
- Architecture Métier
  - Acteurs, Rôles, Unités d'organisation, Objets métier, Processus, Capacités...
- Architecture Applicative
  - Applications, Services, fonctions, Flux, Cartographie...
- Architecture Technique
  - Serveurs, Déploiement, Frameworks, SGBD, Réseau, Cloud...



Figure 20 Gestion des connaissances du SI<sup>11</sup>

<sup>11</sup> Source : <https://www.privaciz.com/architecture-dentreprise/>

## 4 LA DEMARCHE MAGISTHER

La démarche MAGISTHER (**M**éthode et **A**rchitecture de **G**ouvernance de **L'**Information **S**upport Au **T**hermique à Flamme) est un projet de centralisation et exploitation des données afin de valoriser et cartographier notre système d'information. Cette démarche est en développement au sein d'INGEUM depuis 2017. Ce projet s'inscrit dans l'optique de gérer les données comme un actif stratégique de l'entreprise.

### 4.1 CONCEPT

MAGISTHER a pour but de systématiser la construction d'**une cartographie du patrimoine informationnel en continu** actuellement sur un périmètre allant de la capture de la donnée de process industriel, typiquement la donnée issue de capteurs, à sa valorisation. La figure suivante illustre de manière macroscopique la démarche dans son ensemble.

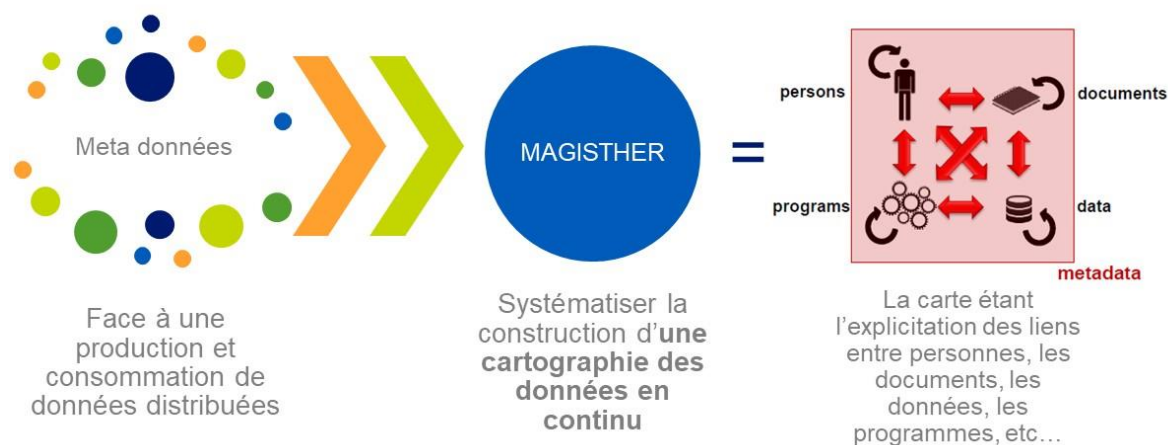


Figure 21 Chaîne de valeur inhérente à MAGISTHER

La chaîne de valeur inhérente à MAGISTHER débute par le fait de capturer des métadonnées à tous les niveaux de l'infrastructure de données. Ces métadonnées doivent permettre de reconstruire une représentation du SI. La capture de ses métadonnées peut passer par le déploiement de Script qui permet l'extraction de nombreux fichiers pouvant être des configurations de logiciels, des modèles de base de données, des inventaires de serveurs, etc... Cela peut également nécessiter d'avoir recours à un acteur humain par une collecte ou saisie manuelle si aucun moyen automatique ne peut être mis en œuvre.

Les scripts sont symbolisés dans la figure suivante par des Robots alors que le responsable humain est symbolisé par le pictogramme de « Responsable ».



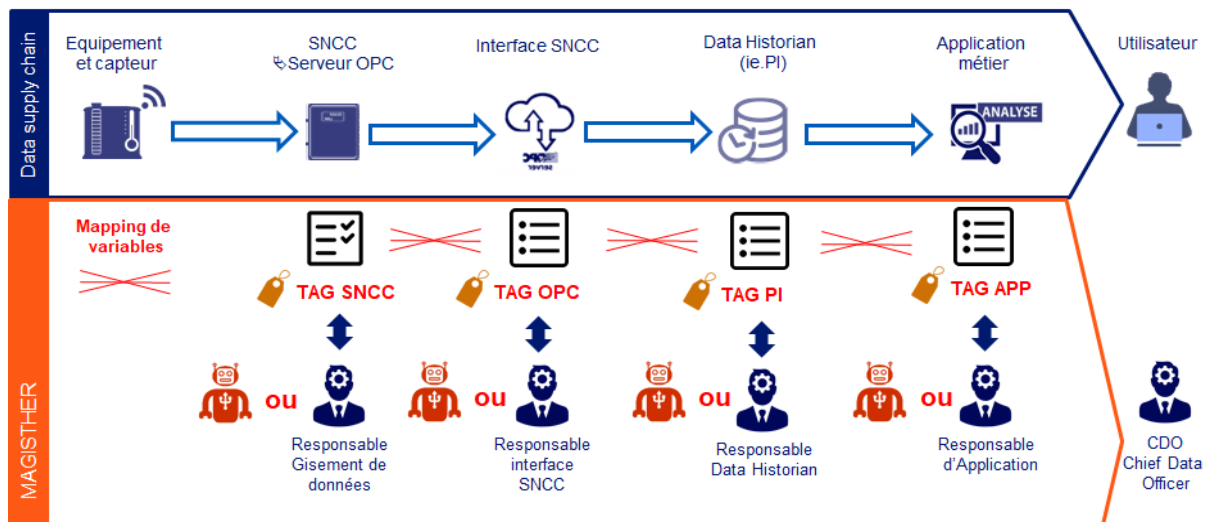


Figure 22 Magister et la chaîne d’approvisionnement

En rapprochant les métadonnées collectées au niveau de différents composants, il est alors possible de reconstituer une représentation de l’infrastructure de données qui constituent alors notre cartographie. Cette cartographie est implémentée sous forme de graphe de données liées qui permet de décrire la filiation entre de nombreux concepts et objet métier de manière dynamique et agile.

Voici un exemple de TAG qui sont liées car l’un alimente l’autre, mais dont le rapprochement par simple identifiant n’est pas possible car ils sont différents dans leur écriture et/ou leur codage :

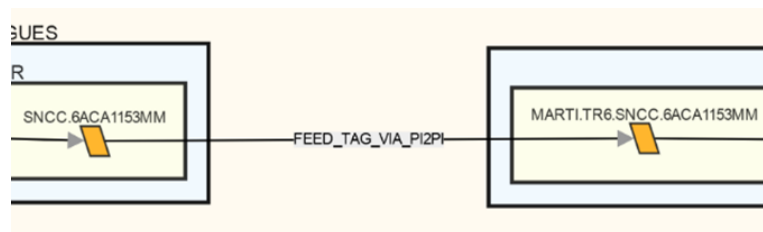


Figure 23 Exemple de lien entre deux identifiants indirect

Ils correspondent à un TAG cloné d’une instance locale (situé à Martigues) à une instance centralisée regroupant l’ensemble des TAG des différents sites.

A la cible, le graphe de connaissance permettra de traduire la filiation entre Données, Information, Personnes, Matériels et au-delà.

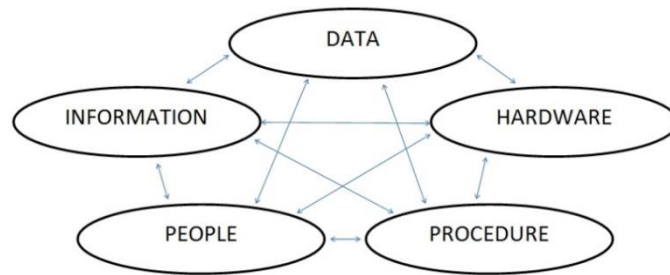


Figure 24 Les composantes du système d'information

Par analogie, on peut dire que si les objets métiers du SI, identifiables via les métadonnées, sont des étoiles MAGISTHER doit permettre de construire une carte du ciel en mettant en évidence les étoiles formant des constellations. Ceci rend alors visible les relations tangibles ou intangibles existantes entre les objets qui composent l'univers numérique inhérent à nos infrastructures SI.

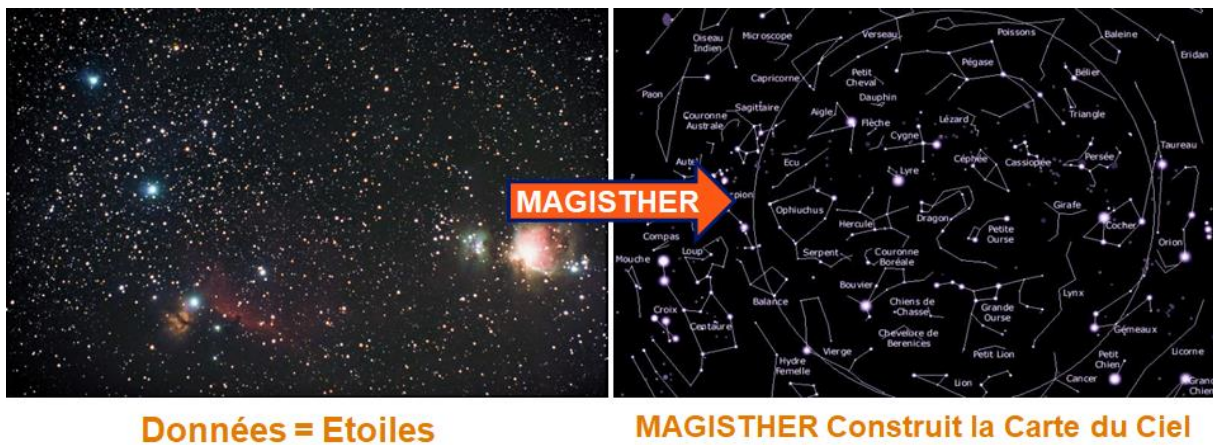


Figure 25 Analogie entre les données et les étoiles

L'approche consiste alors à identifier dans chaque couche du SI, les référentiels stables existants, numérique ou à numériser, ainsi que les objets pouvant être inventoriés dans la mesure du possible automatiquement.

A partir de ce travail d'identification, nous définissons la manière dont l'ensemble des objets peuvent être maillés les uns avec les autres au sein d'une même couche ou entre différentes couches. Il en résulte une indexation du système d'information qui peut s'apparenter à un embryon de jumeau numérique du SI.

A partir de ce graphe de données liées, il est dès lors possible de générer automatiquement des cartographies interactives à destination de l'humain.

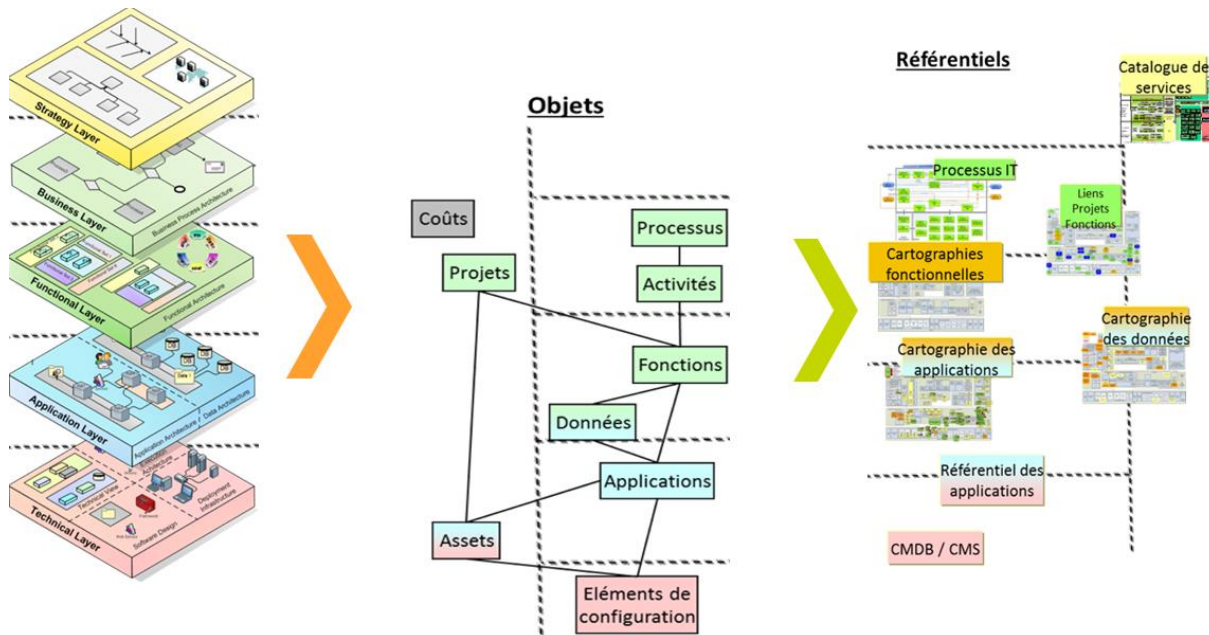


Figure 26 Le SI sous forme de couche caractérise en son sein des objets eux même appartenant à des référentiels

## 4.2 PLATEFORME MAGISTHER

La plateforme MAGISTHER se matérialise par l'intégration de différents outils afin d'instrumenter la démarche de gouvernance. L'objectif premier de cette plateforme numérique est de capturer des données de référence en vue de constituer **un référentiel sémantique**. Ce référentiel sémantique constitue une base de connaissances (les données étant contextualisées) au service d'acteurs humains (M2H : Machine to Human) ou virtuels (M2M : Machine to Machine).

Le fonctionnement de la plateforme numérique MAGISTHER peut être résumé en 4 grandes étapes :

- Capturer des données de référence,
- Historiser les données de référence,
- Construction d'un référentiel sémantique sous la forme de données liées (ie. Linked Data) agnostique des formats propriétaires,
- Exposer des services à partir de ce référentiel sémantique sous la forme d'un portail WEB.

La figure suivante reprend ces 4 étapes de manière schématique.

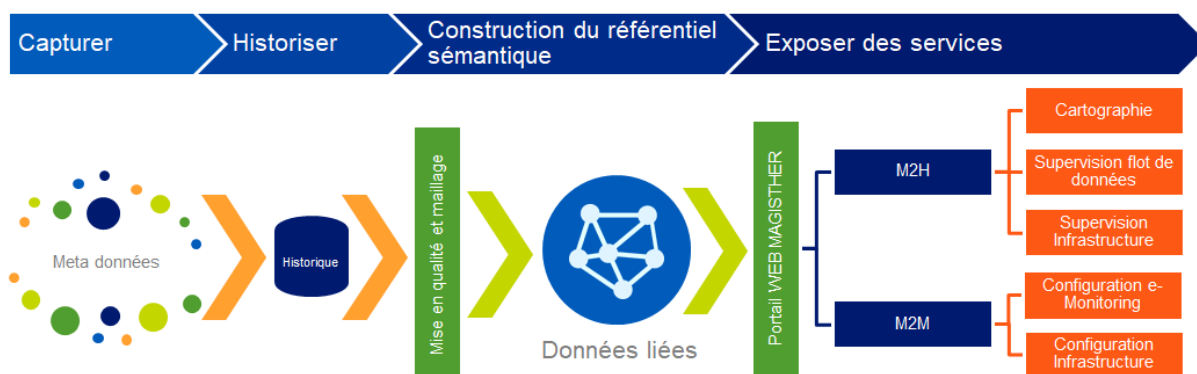


Figure 27 Vision macroscopique du fonctionnement de la plateforme MAGISTHER

La **capture des données de référence privilégie au maximum la collecte automatique de données** numériques par extraction à partir de sources de données existantes jouissant d'une certaine pérennité et affichant des formats d'export stables. Le référentiel sémantique visant un spectre large d'usages, une partie de celui-ci pourrait faire l'objet d'une alimentation par des utilisateurs mais via des interfaces de saisie contrôlées.

L'ensemble des données qui contribue à la fabrication du référentiel fait l'objet d'une **Historisation** (sauvegarde et d'un « versionning ») avant même que celui-ci soit calculé. Nous nous appuyons ici sur la technologie de versionning de code logiciel Git<sup>12</sup>.

La **construction du référentiel sémantique** est un processus automatique s'exécutant à chaque fois qu'une modification des données de référence est détectée. Le calcul du référentiel peut également être invoqué si la stratégie de maillage des données de référence (ie. L'architecture du graphe de connaissances) évolue.

L'exposition de services à partir du référentiel cible deux orientations principales :

- A destination de l'humain (M2H) en exposant des services d'IHM :
  - Afin de simplifier la manière d'accéder à des données de référence ;
  - Afin d'avoir une meilleure compréhension des données utilisées ;
  - Afin d'avoir une meilleure compréhension du système d'information dans sa globalité.
- A destination de machines (M2M) :
  - Afin de générer des fichiers à partir du référentiel ;
  - Afin d'exposer des services (ie. API) de données.

A titre d'exemple, nous pouvons décrire le module de **catalogue de données** des sites.

Ce module permet d'accéder à l'ensemble des données récoltées sur chacun des sites pris en charge par la DTEAM (figure 25). Ces catalogues intègrent des fonctions de filtrage et de regroupement par concept rendant la navigation fluide et plus rapide. Toutes les données présentées dans les catalogues sont le résultat d'une requête dans notre pivot sémantique que

<sup>12</sup> Source : <https://en.wikipedia.org/wiki/Git>

constitue notre base de graphes. Ainsi, en ajoutant des données, des attributs ou des concepts dans notre cartographie, ils seront automatiquement retranscrits dans les différents modules.

| code_system | pointsource | codification_type    | description | engunits               | code_equipement | label | engunits_label         |
|-------------|-------------|----------------------|-------------|------------------------|-----------------|-------|------------------------|
| ACV...      | KIL         | /CETAC/POINT_SOUR... | ECS         | ALM_CHAINE_TECHN...    | /CETAC/UOM/     | TC    | TAC.VA3.TC.1K1L0090... |
| ACV...      | GEV         | /CETAC/POINT_SOUR... | ECS         | POSITION DISCORDA...   | /CETAC/UOM/     | JA    | TAC.VA3.TC.3GEV000...  |
| ACV...      | GEV         | /CETAC/POINT_SOUR... | ECS         | POSITION INTERMEDI...  | /CETAC/UOM/     | JA    | TAC.VA3.TC.3GEV000...  |
| ACV...      | GEV         | /CETAC/POINT_SOUR... | ECS         | POSITION DISONCTE...   | /CETAC/UOM/     | JA    | TAC.VA3.TC.3GEV000...  |
| ACV...      | GEV         | /CETAC/POINT_SOUR... | ECS         | POSITION DISONCTE...   | /CETAC/UOM/     | JA    | TAC.VA3.TC.3GEV000...  |
| ACV...      | GEV         | /CETAC/POINT_SOUR... | ECS         | POSITION DISCORDA...   | /CETAC/UOM/     | JA    | TAC.VA3.TC.3GEV000...  |
| ACV...      | GEV         | /CETAC/POINT_SOUR... | ECS         | POSITION INTERMEDI...  | /CETAC/UOM/     | JA    | TAC.VA3.TC.3GEV000...  |
| ACV...      | GEV         | /CETAC/POINT_SOUR... | ECS         | POSITION DISONCTE...   | /CETAC/UOM/     | JA    | TAC.VA3.TC.3GEV000...  |
| ACV...      | GEV         | /CETAC/POINT_SOUR... | ECS         | POSITION DISONCTE...   | /CETAC/UOM/     | JA    | TAC.VA3.TC.3GEV000...  |
| ACV...      | -           | /CETAC/POINT_SOUR... | -           | VITESSE                | /CETAC/UOM/rpm  | -     | TAC.VA3.TC.3GRE000...  |
| ACV...      | -           | /CETAC/POINT_SOUR... | -           | MODE FULL SPEED N...   | /CETAC/UOM/     | -     | TAC.VA3.TC.3GRE000...  |
| ACV...      | -           | /CETAC/POINT_SOUR... | -           | ALARME FEU TAC         | /CETAC/UOM/     | -     | TAC.VA3.TC.3JDT0001... |
| ACV...      | -           | /CETAC/POINT_SOUR... | -           | DEFAULT CENTRALE IN... | /CETAC/UOM/     | -     | TAC.VA3.TC.3JDT0002... |
| ACV...      | -           | /CETAC/POINT_SOUR... | -           | DEF ALIM ELEC CENT...  | /CETAC/UOM/     | -     | TAC.VA3.TC.3JDT0003... |
| ACV...      | KCO         | /CETAC/POINT_SOUR... | ECS         | Mesure PA              | /CETAC/UOM/MW   | MW    | TAC.VA3.TC.3KCO000...  |
| ACV...      | KCO         | /CETAC/POINT_SOUR... | ECS         | Mesure Q               | /CETAC/UOM/MVAh | MQ    | TAC.VA3.TC.3KCO000...  |
| ACV...      | -           | /CETAC/POINT_SOUR... | -           | DEFAULT MINEUR         | /CETAC/UOM/     | -     | TAC.VA3.TC.3KCO001...  |
| ACV...      | -           | /CETAC/POINT_SOUR... | -           | DEFAULT MAJEUR         | /CETAC/UOM/     | -     | TAC.VA3.TC.3KCO001...  |
| ACV...      | KCO         | /CETAC/POINT_SOUR... | ECS         | MAX ATTEINT            | /CETAC/UOM/     | TC    | TAC.VA3.TC.3KCO009...  |
| ACV...      | -           | /CETAC/POINT_SOUR... | -           | INDISPONIBLE           | /CETAC/UOM/     | -     | TAC.VA3.TC.3KCO200...  |
| ACV...      | -           | /CETAC/POINT_SOUR... | -           | DEMARRAGE TAC          | /CETAC/UOM/     | -     | TAC.VA3.TC.3KCO200...  |
| ACV...      | -           | /CETAC/POINT_SOUR... | -           | TRIP                   | /CETAC/UOM/     | -     | TAC.VA3.TC.3KCO200...  |
| ACV...      | KCO         | /CETAC/POINT_SOUR... | ECS         | MESURE PMD             | /CETAC/UOM/MW   | MW    | TAC.VA3.TC.3KCO204...  |
| ACV...      | KCO         | /CETAC/POINT_SOUR... | ECS         | MT                     | /CETAC/UOM/MW   | MW    | TAC.VA3.TC.3KCO204...  |
| ACV...      | KCO         | /CETAC/POINT_SOUR... | ECS         | PCMAX                  | /CETAC/UOM/MW   | MW    | TAC.VA3.TC.3KCO204...  |

Figure 28 Module d'accès aux catalogues de données des sites

| pointsource_label | uid                   | code_system | pointsource          | codification_type | description            | engunits        | code_equipement | label                  | engunits_label |
|-------------------|-----------------------|-------------|----------------------|-------------------|------------------------|-----------------|-----------------|------------------------|----------------|
| OPC-TC            | /CETAC/TAG_P/ITACV... | KIL         | /CETAC/POINT_SOUR... | ECS               | ALM_CHAINE_TECHN...    | /CETAC/UOM/     | TC              | TAC.VA3.TC.1K1L0090... |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | GEV         | /CETAC/POINT_SOUR... | ECS               | POSITION DISCORDA...   | /CETAC/UOM/     | JA              | TAC.VA3.TC.3GEV000...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | GEV         | /CETAC/POINT_SOUR... | ECS               | POSITION INTERMEDI...  | /CETAC/UOM/     | JA              | TAC.VA3.TC.3GEV000...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | GEV         | /CETAC/POINT_SOUR... | ECS               | POSITION DISONCTE...   | /CETAC/UOM/     | JA              | TAC.VA3.TC.3GEV000...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | GEV         | /CETAC/POINT_SOUR... | ECS               | POSITION DISONCTE...   | /CETAC/UOM/     | JA              | TAC.VA3.TC.3GEV000...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | GEV         | /CETAC/POINT_SOUR... | ECS               | POSITION DISCORDA...   | /CETAC/UOM/     | JA              | TAC.VA3.TC.3GEV000...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | GEV         | /CETAC/POINT_SOUR... | ECS               | POSITION INTERMEDI...  | /CETAC/UOM/     | JA              | TAC.VA3.TC.3GEV000...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | GEV         | /CETAC/POINT_SOUR... | ECS               | POSITION DISONCTE...   | /CETAC/UOM/     | JA              | TAC.VA3.TC.3GEV000...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | GEV         | /CETAC/POINT_SOUR... | ECS               | POSITION DISONCTE...   | /CETAC/UOM/     | JA              | TAC.VA3.TC.3GEV000...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | -           | /CETAC/POINT_SOUR... | -                 | VITESSE                | /CETAC/UOM/rpm  | -               | TAC.VA3.TC.3GRE000...  | rpm            |
| OPC-TC            | /CETAC/TAG_P/ITACV... | -           | /CETAC/POINT_SOUR... | -                 | MODE FULL SPEED N...   | /CETAC/UOM/     | -               | TAC.VA3.TC.3GRE000...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | -           | /CETAC/POINT_SOUR... | -                 | ALARME FEU TAC         | /CETAC/UOM/     | -               | TAC.VA3.TC.3JDT0001... |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | -           | /CETAC/POINT_SOUR... | -                 | DEFAULT CENTRALE IN... | /CETAC/UOM/     | -               | TAC.VA3.TC.3JDT0002... |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | -           | /CETAC/POINT_SOUR... | -                 | DEF ALIM ELEC CENT...  | /CETAC/UOM/     | -               | TAC.VA3.TC.3JDT0003... |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | KCO         | /CETAC/POINT_SOUR... | ECS               | Mesure PA              | /CETAC/UOM/MW   | MW              | TAC.VA3.TC.3KCO000...  | MW             |
| OPC-TC            | /CETAC/TAG_P/ITACV... | KCO         | /CETAC/POINT_SOUR... | ECS               | Mesure Q               | /CETAC/UOM/MVAh | MQ              | TAC.VA3.TC.3KCO000...  | MVAh           |
| OPC-TC            | /CETAC/TAG_P/ITACV... | -           | /CETAC/POINT_SOUR... | -                 | DEFAULT MINEUR         | /CETAC/UOM/     | -               | TAC.VA3.TC.3KCO001...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | -           | /CETAC/POINT_SOUR... | -                 | DEFAULT MAJEUR         | /CETAC/UOM/     | -               | TAC.VA3.TC.3KCO001...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | KCO         | /CETAC/POINT_SOUR... | ECS               | MAX ATTEINT            | /CETAC/UOM/     | TC              | TAC.VA3.TC.3KCO009...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | -           | /CETAC/POINT_SOUR... | -                 | INDISPONIBLE           | /CETAC/UOM/     | -               | TAC.VA3.TC.3KCO200...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | -           | /CETAC/POINT_SOUR... | -                 | DEMARRAGE TAC          | /CETAC/UOM/     | -               | TAC.VA3.TC.3KCO200...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | -           | /CETAC/POINT_SOUR... | -                 | TRIP                   | /CETAC/UOM/     | -               | TAC.VA3.TC.3KCO200...  |                |
| OPC-TC            | /CETAC/TAG_P/ITACV... | KCO         | /CETAC/POINT_SOUR... | ECS               | MESURE PMD             | /CETAC/UOM/MW   | MW              | TAC.VA3.TC.3KCO204...  | MW             |
| OPC-TC            | /CETAC/TAG_P/ITACV... | KCO         | /CETAC/POINT_SOUR... | ECS               | MT                     | /CETAC/UOM/MW   | MW              | TAC.VA3.TC.3KCO204...  | MW             |
| OPC-TC            | /CETAC/TAG_P/ITACV... | KCO         | /CETAC/POINT_SOUR... | ECS               | PCMAX                  | /CETAC/UOM/MW   | MW              | TAC.VA3.TC.3KCO204...  | MW             |

Figure 29 Module d'accès aux catalogues ciblant un site : CETAC

Nous pouvons également aborder le module de **Cartographie**.

Ce module correspond à la représentation la plus proche de notre pivot sémantique. Il permet de représenter les données contenues dans la base de graphes sous forme de nœuds et de relations. Nous pouvons ainsi accéder à l'ensemble des informations rattachées à un nœud (avec l'ouverture d'un panel figure 28). Nous pouvons également envoyer des requêtes sur le graphe afin d'obtenir le voisinage (ou le lignage) d'un nœud. Cette requête peut intégrer un filtrage et ainsi limiter les résultats à une couche précise (donnée, infrastructure...) Dans l'exemple de la figure 30, nous avons demandé le lignage complet, limité à la couche "Data" du point PI affiché

en surbrillance. Cette requête nous permet d'obtenir les points consommés (ou qui consomme) le point cible. Dans notre exemple, nous constatons qu'au sein de la Stack Monitor, le point cible est alimenté par un second (1). Ce dernier est lui-même alimenté dans une autre infrastructure (Stack Martigues) ... Nous reconnaissons bien la reconstitution de la chaîne de collecte des informations citées page 27.

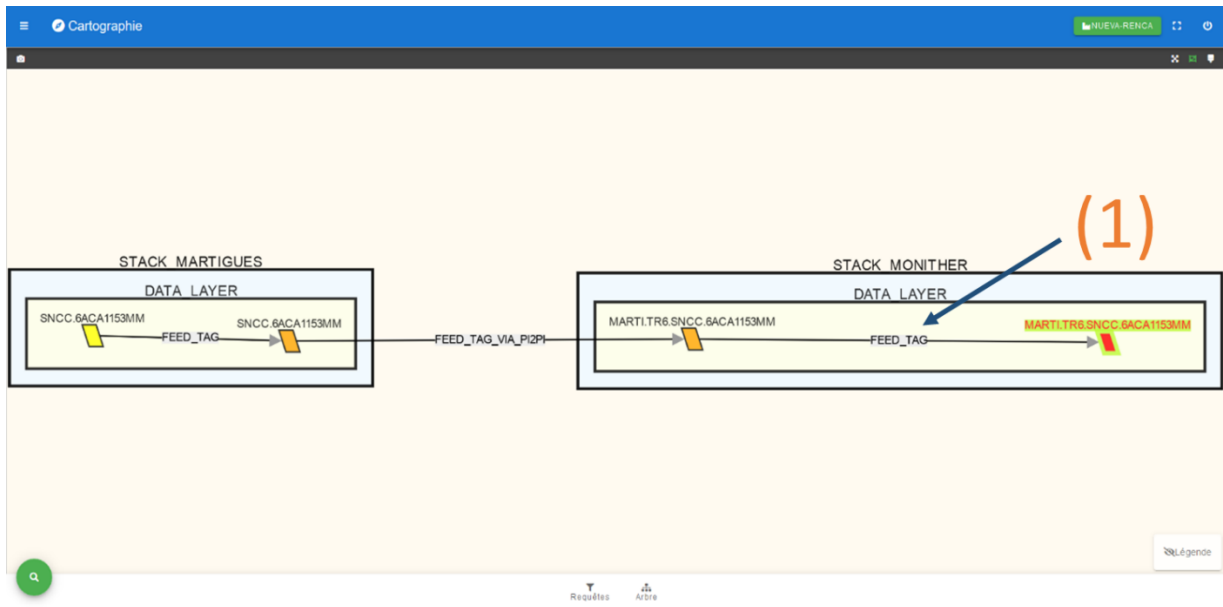


Figure 30 Module de cartographie illustrant les données liées

| Attribut          | Valeur                                    |
|-------------------|---|
| cia               |   |
| pointsource_label | PIZPI-MARTIGUES                           |
| uid               | /MONITHER/TAG_PIMARTI.TR6.SNCC.6ACA1153MM |
| code_system       | -   |
| pointsource       | /MONITHER/POINT_SOURCE/PIZPI-MARTIGUES    |
| codification_type | -   |
| description       | REGL F ONTOUR ECON BP POS                 |
| code_equipement   | -   |
| engunits          | /MONITHER/UOM/%                           |
| engunits_label    | %   |
| label             | MARTI.TR6.SNCC.6ACA1153MM                 |

Figure 31 Module de cartographie présentant les données d'un nœud

## 5 MAGISTHER LE PROJET FIL ROUGE DE MON ALTERNANCE

Le développement et le déploiement de la démarche MAGISTHER ont constitué un réel fil rouge tout au long des 3 ans de mon alternance. J'ai ainsi suivi l'ensemble de la démarche et j'ai contribué de manière plus ou moins importante à chaque étage de la solution.

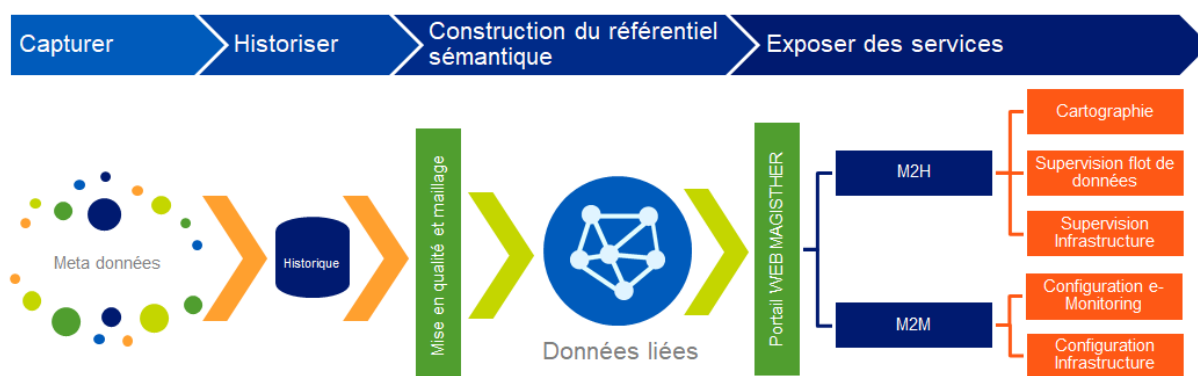


Figure 32 Vision macroscopique du fonctionnement de la plateforme MAGISTHER

### 5.1 PILOTAGE DE LA CAPTURE DES METADONNEES

La première étape du développement de notre solution consiste à capturer et historiser les métadonnées utiles. Notre objectif est double :

- Identifier les métadonnées nécessaires à la génération de notre cartographie,
- Historiser (ie versionner) ces métadonnées afin de conserver une trace de toute modification de ces dernières,
- Générer notre cartographie à l'aide des données capturées.

Dans un premier temps, nous avons identifié les métadonnées nécessaires à l'établissement de la cartographie et dû identifier le moyen de les capturer. A partir de là, nous avons pu concevoir l'implémentation nécessaire.

Pour mettre en œuvre la capture des métadonnées et leur historisation, j'ai coordonné l'activité de deux prestataires experts dans ces domaines. Le travail a consisté à faire écrire des scripts de récupération de ces métadonnées, de les déployer dans l'infrastructure SI et d'orchestrer leur invocation. La mécanique mise en place permet la collecte automatique des métadonnées à partir de leurs sources respectives pour les historiser.

Nous avons pour cela mis en œuvre un dépôt Git au sein d'une plateforme Gitlab. Gitlab une forge logicielle permettant à plusieurs développeurs de collaborer sur un même projet informatique. Il permet de répertorier l'intégralité des fichiers qui composent un projet sous forme de dépôt. Il intègre un accès collaboratif à chaque fichier stocké et une historisation des modifications apportées à ces derniers. Il permet également d'exécuter des scripts de façon

automatique. La plateforme Gitlab est une solution mutualisée qui n'est pas dédiée à la démarche MAGISTHER.

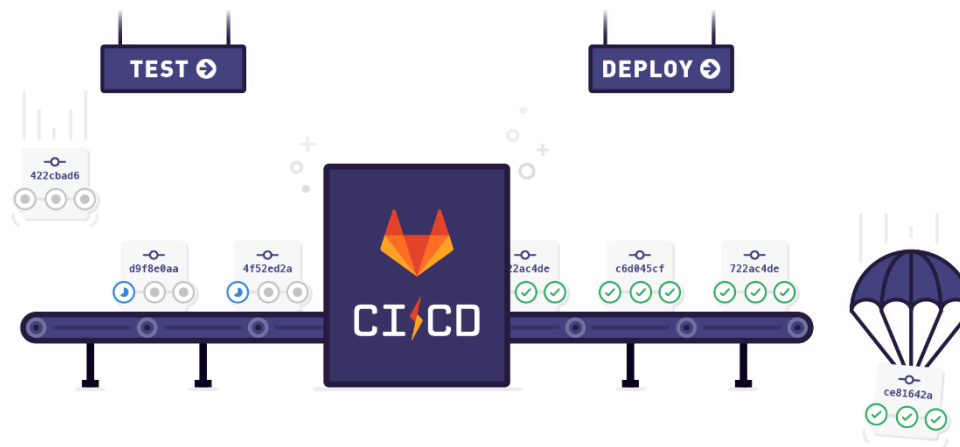


Figure 33 Illustration du fonctionnement de GITLAB<sup>13</sup>

Une fois les données copiées au sein du dépôt, notre ingénieur système a configuré les différentes routines d'exécution de script correspondant à l'orchestration nécessaire pour collecter nos métadonnées. Les scripts de récupération ont été automatisés pour que tous les matins, une mise à jour des données soit effectuée.

## 5.2 LE PRE TRAITEMENT DES METADONNEES VIA LA CONCEPTION DE CHAINES ETL

Parmi mes contributions au projet MAGISTHER je suis intervenu dans l'étage ETL.

En effet, pour construire notre base de connaissances, nous avons besoin d'information.

Ces informations sont des données de référence (1) ainsi que des données et métadonnées métier (2) comme illustré sur la figure suivante (figure 34). Mais ces informations ne sont pas liées entre elles et possèdent une structure fixée par leur source d'origine (structure d'un fichier variant d'une source à l'autre).

Or, pour que notre base de données soit pertinente, il faut que nos données soient uniformisées. Si chaque information recherchée ne peut être rapproché d'une autre c'est que notre gouvernance se révèle peu précise.

Pour ce faire, nous exploitons des outils de transformation de l'information telle que Talend<sup>14</sup> ou des langages de transformation de données comme le XSL. Ces scripts nous permettent de construire des fichiers adéquats pour la génération de notre base de connaissances et permettant de construire notre cartographie.

<sup>13</sup> Source : <https://medium.com/faun/building-a-docker-image-with-gitlab-ci-and-net-core-8f59681a86c4>

<sup>14</sup> Source : <https://fr.talend.com/>



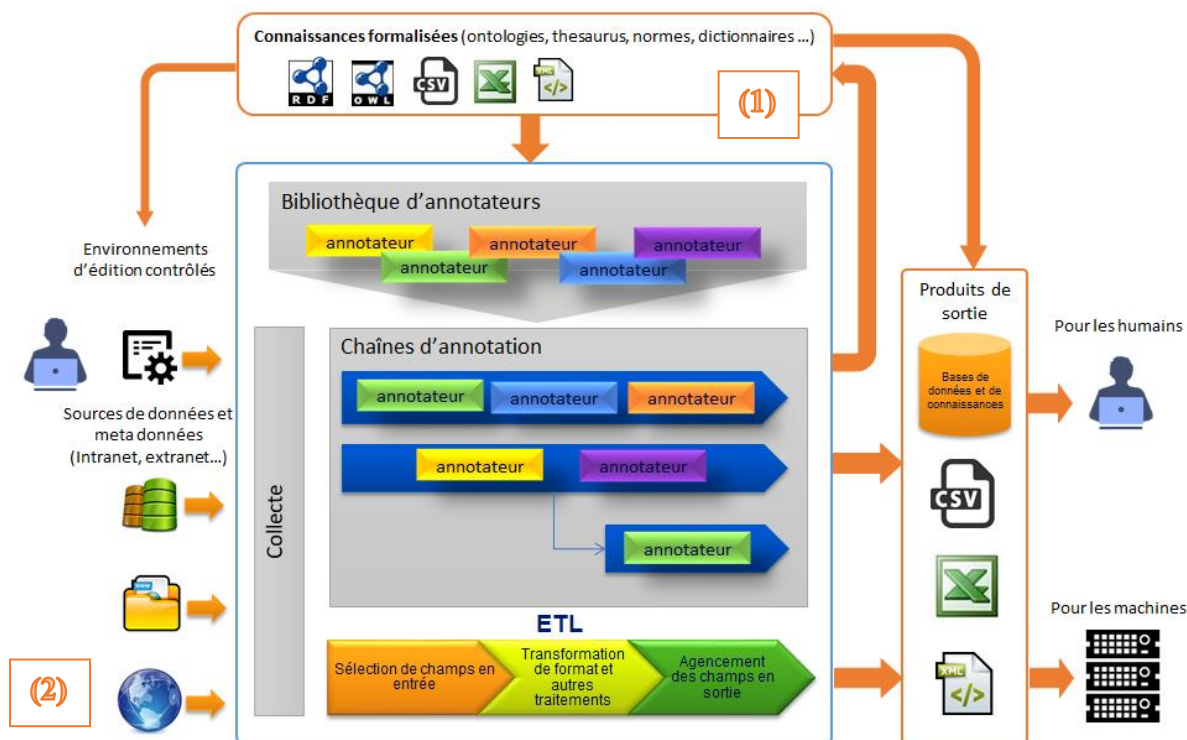


Figure 34 Fonctionnement interne de MAGISTHER

### 5.3 LA CARTOGRAPHIE SOUS LA FORME D'UN GRAPHE DE CONNAISSANCES

Afin de construire notre base de connaissances et donc notre cartographie, nous exploitons une base de données orientée graphe.

Ce type de base de données nous permet de stocker des données liées et les relations entre ses données. Grâce à cette caractéristique, cette technologie est particulièrement adaptée à la construction de référentielles sémantiques qui se traduisent sous la forme de modèles complexes.

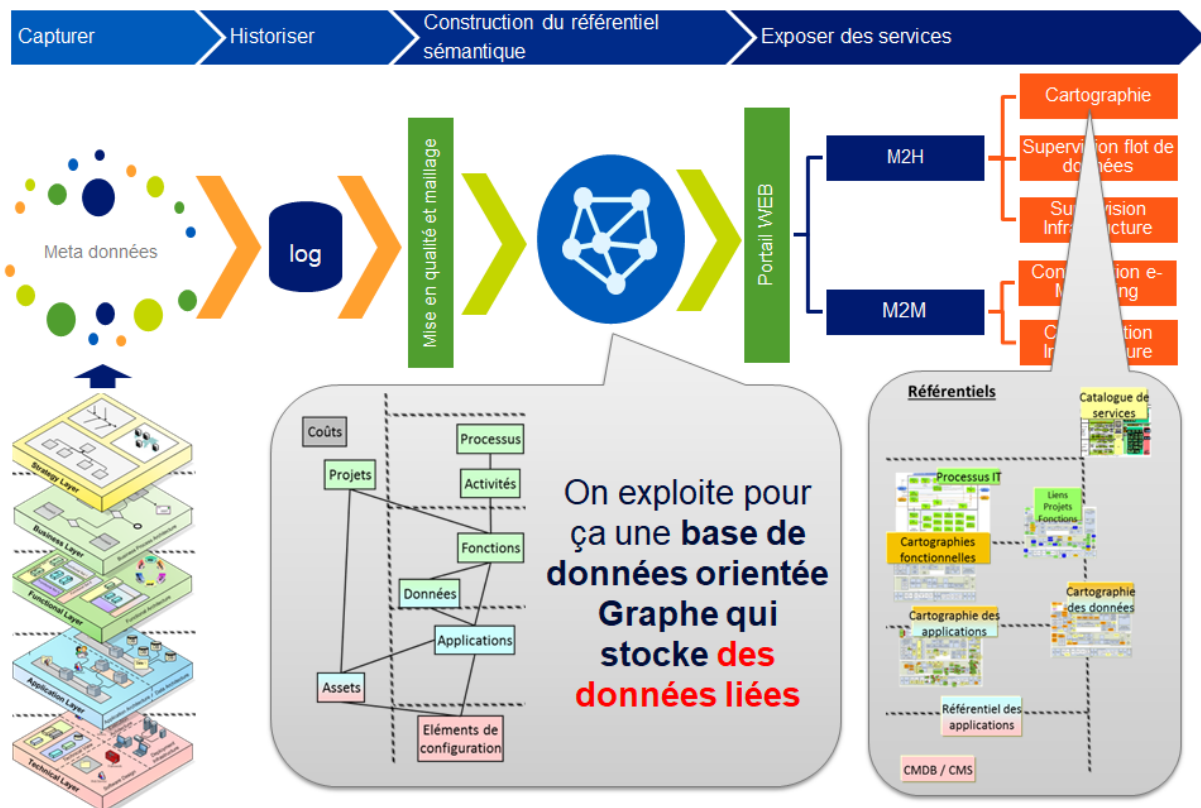


Figure 35 Vision macroscopique du fonctionnement de la plateforme MAGISTHER (détaillée)

Ce type de base de données nous permet de stocker des données liées au sens où des relations sont créées entre ses données.

Les bases de données orientées graphe sont pilotables via à un langage de requête spécifique pour cette technologie : Le Cypher<sup>15</sup>.

Ce langage permet de structurer les informations dans notre base de données (tracer les relations, créer les nœuds, et associer des attributs). Ce langage permet également de requêter la base une fois cette dernière construite.

<sup>15</sup> Source : [https://fr.wikipedia.org/wiki/Cypher\\_\(langage\)](https://fr.wikipedia.org/wiki/Cypher_(langage))

Dans le système « PAC », je voudrais tous les points de mesure utilisés par un projet de DPA Prism

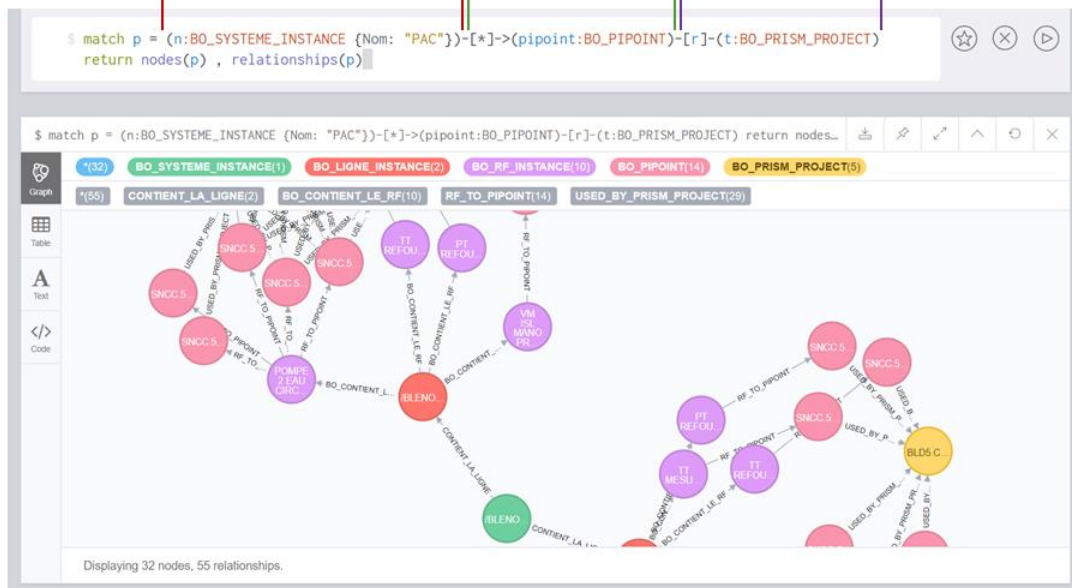


Figure 36 Exemple d'utilisation des requêtes Cypher

Je suis donc intervenu dans la construction des fichiers Cypher.

Le Cypher exploite les fichiers générés à l'aide de notre ETL. En fonction de ces fichiers sources, j'ai écrit un fichier décrivant chaque nœud avec les différents attributs associés à ces derniers, ainsi que les relations entre chacun de ces nœuds. Tous ces éléments sont ajoutés au script pour chaque infrastructure et décrivent les différentes couches du SI (symbolisées par les groupements jaunes dans les figures suivantes). Ce script définit la structure de notre base de connaissances.

A l'heure actuelle, nous avons la cartographie de l'ensemble des données et les relations entre elles.

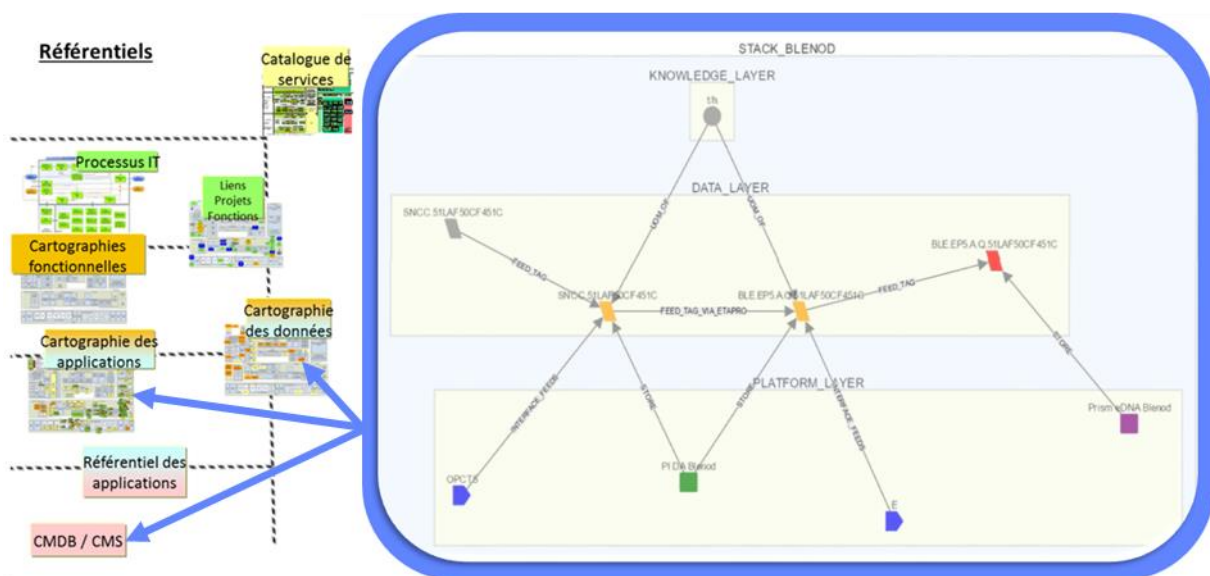


Figure 37 Intégration de la cartographie des données dans notre portail

Nous avons construit une cartographie décrivant 3 couches du SI (couches techniques, applicative et fonctionnelle).

Nous avons donc une cartographie sur laquelle nous pouvons requêter à travers un portail web (développé par notre équipe). Le portail permet d'obtenir le lignage complet d'un nœud (représentant n'importe quelle donnée présente dans le graphe).

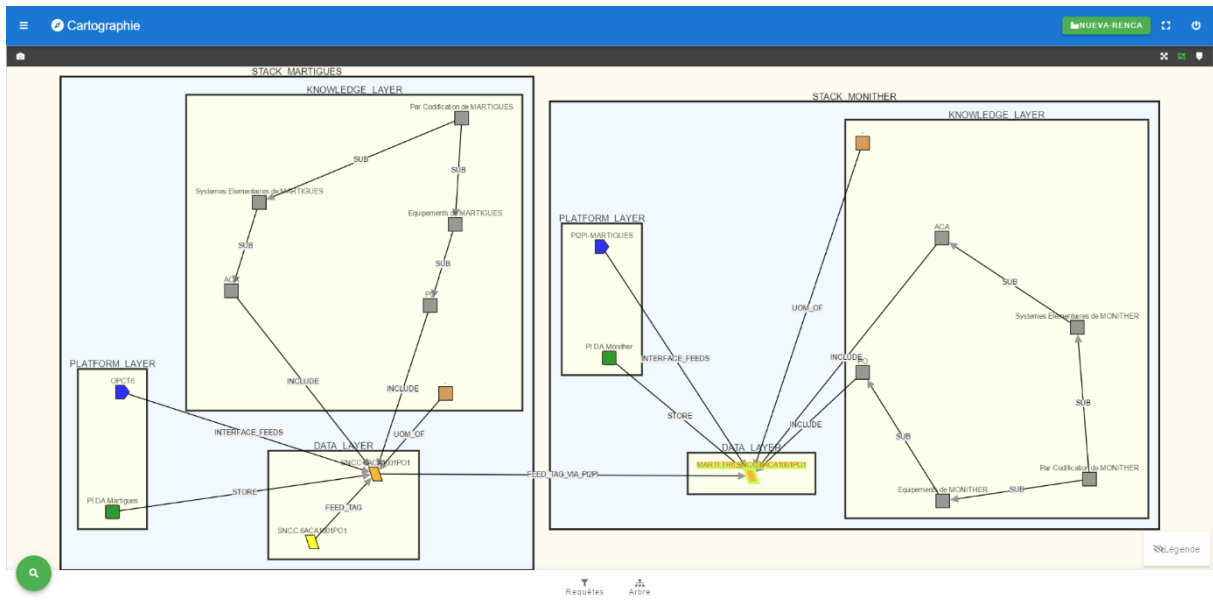


Figure 38 Exemple de représentation des données dans le module de la cartographie

Cela nous permet de répondre aux questions suivantes :

Quel est le cheminement d'une mesure au travers plusieurs infrastructures (les carrés bleus représentant les différentes infrastructures). Ainsi nous répondons à la problématique de reconstitution de chaîne d'approvisionnement des données. Cela nous permet également de constater indirectement les impacts d'une modification (ou suppression) d'un point sur l'ensemble de son voisinage.

Sur la figure suivante, nous constatons que notre TAG en surbrillance est alimenté par un second dans une infrastructure différente via un lien PtoPI.

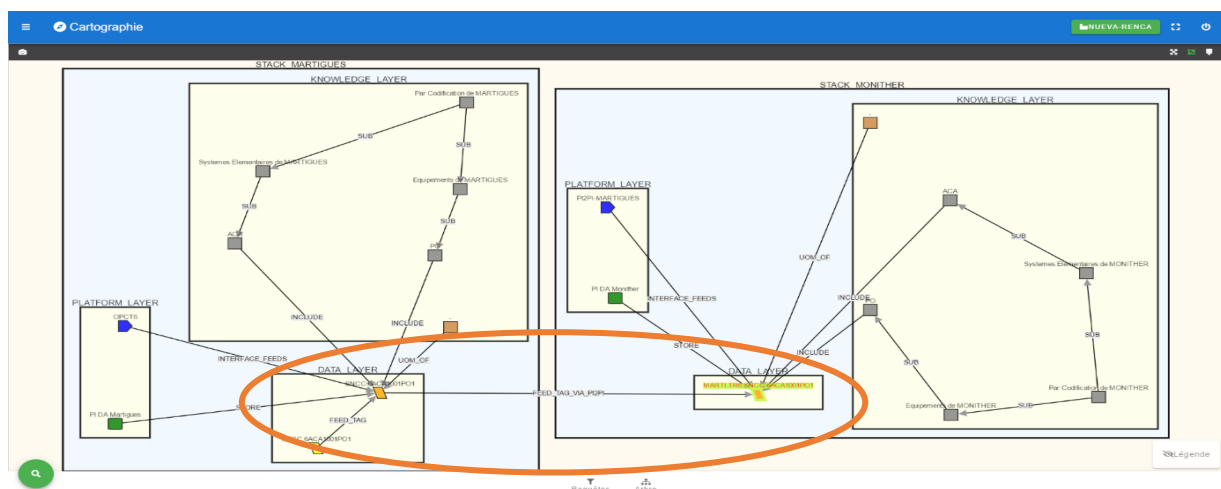


Figure 39 Exemple de lien entre différentes instances de TAG

Nous retrouvons également des informations à propos des sources de la donnée ciblée. Dans l'exemple ci-dessous nous pouvons voir dans la couche Plateforme qu'il y a une interface de Martigues qui symbolise la source d'alimentation de la donnée en surbrillance.

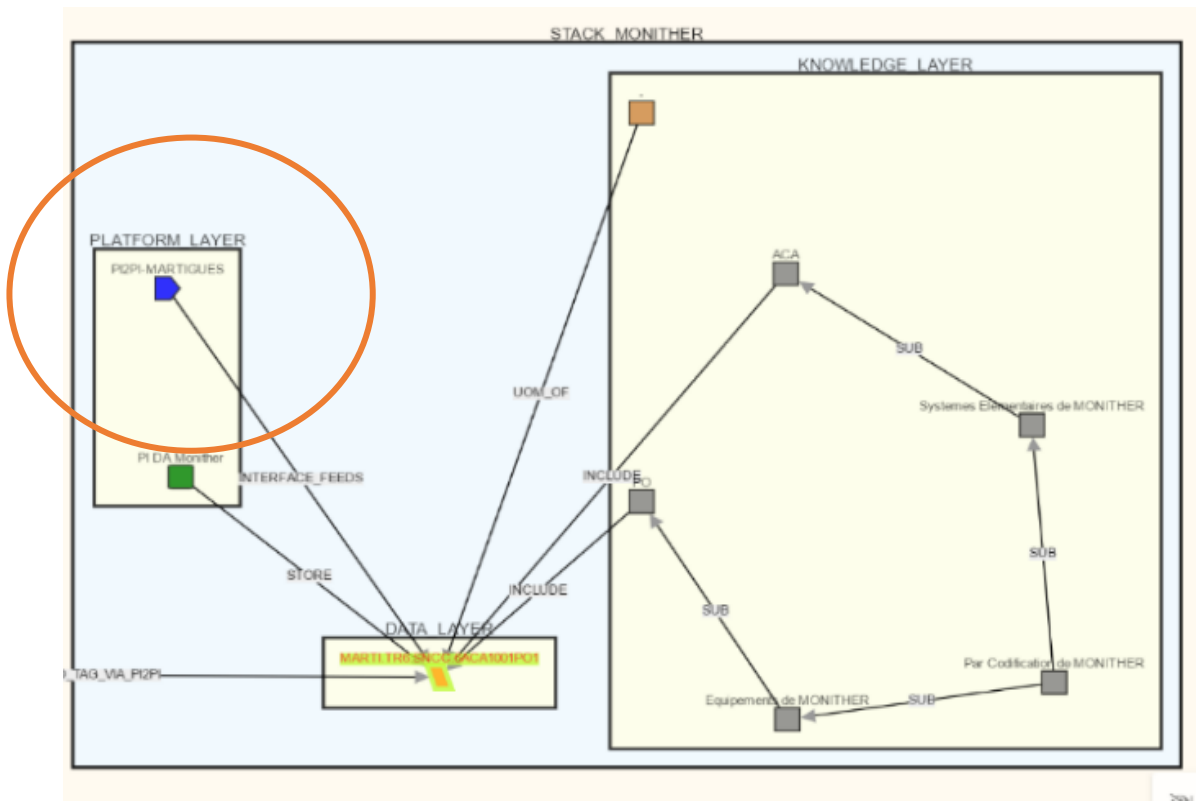


Figure 40 Exemple de lien avec la source du TAG

Avec de telles fonctionnalités, nous répondons aux objectifs de la plateforme MAGISTER se traduisant par la reconstruction de la chaîne d'approvisionnement en incorporant le mapping des variables et les connexions entre chaque instance des TAGs de la chaîne. En intégrant les interfaces et les sources des différents points, nous pourrions par la suite rattacher un responsable à chaque étape du transport et ainsi résoudre les problèmes d'acheminement en d'adressant à la bonne personne sans intermédiaire.

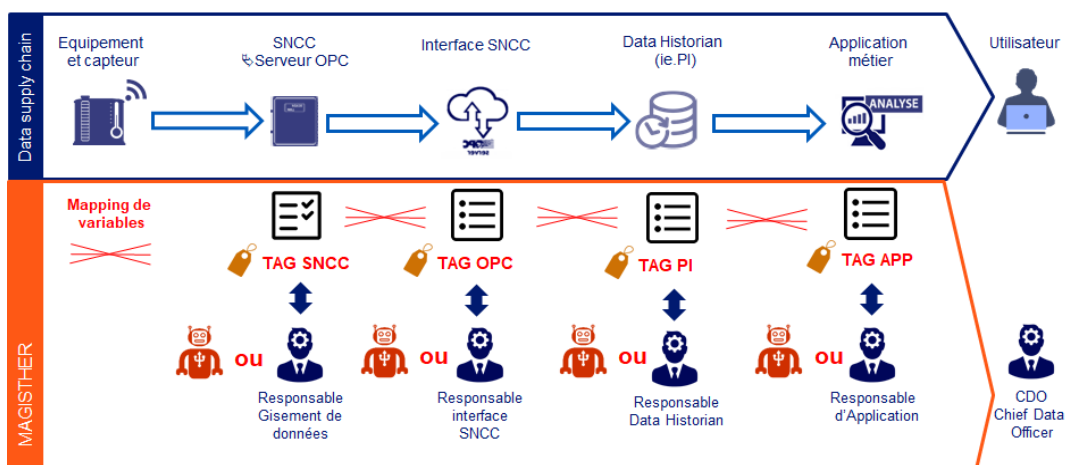


Figure 41 Magister et la chaîne d'approvisionnement

## 6 DE LA GOUVERNANCE DE LA DONNEE A LA SUPERVISION DU SI : EXTENSION DU PERIMETRE DE MAGISTHER

Au sein de MAGISTHER, nous nous appuyons sur une base de graphes qui traduit la cartographie des données. Nous avons initié un élargissement du périmètre afin de décrire plus largement le SI.

Cette extension se pose selon trois axes :

- La génération d'interfaces de supervision automatique en fonction de la configuration (englobant la couche matérielle du modèle) ;
- L'indexation de la documentation directement dans le graphe ;
- La génération automatique de la documentation en fonction des données du graphe.

Ces trois nouveaux chantiers se situent dans la famille de service M2H.

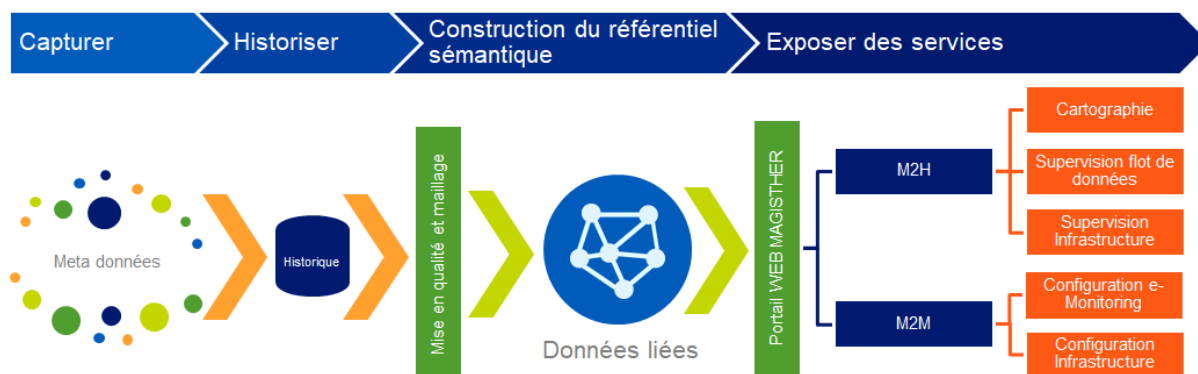


Figure 42 Vision macroscopique du fonctionnement de la plateforme MAGISTHER

La génération d'interfaces de supervision vise à **exploiter les données contenues dans le graphe** pour générer des tableaux de surveillance personnalisée de façon automatique. En incorporant les informations d'architecture réseaux au sein de notre base de connaissances, nous serons en mesure d'exploiter ses informations afin de générer des écrans de supervision s'appuyant sur le jumeau numérique de notre SI.

En ce qui concerne l'indexation de la documentation dans le graphe, nous allons **rattacher des documents à des nœuds du graphe** sur la base de concept commun. Cela permettra, en un seul « click » sur un nœud du graphe, l'ensemble des documents rattachés à ce nœud. Nous aurons donc un accès rapide à un panel d'information sur un concept choisi dans le graphe.

Enfin, la génération automatique de documentation consiste à **exploiter les données présentes dans le graphe** afin de créer de la documentation à destination de l'humain. Cela permettra une pérennisation, une uniformité et une mise à jour constante de la documentation.

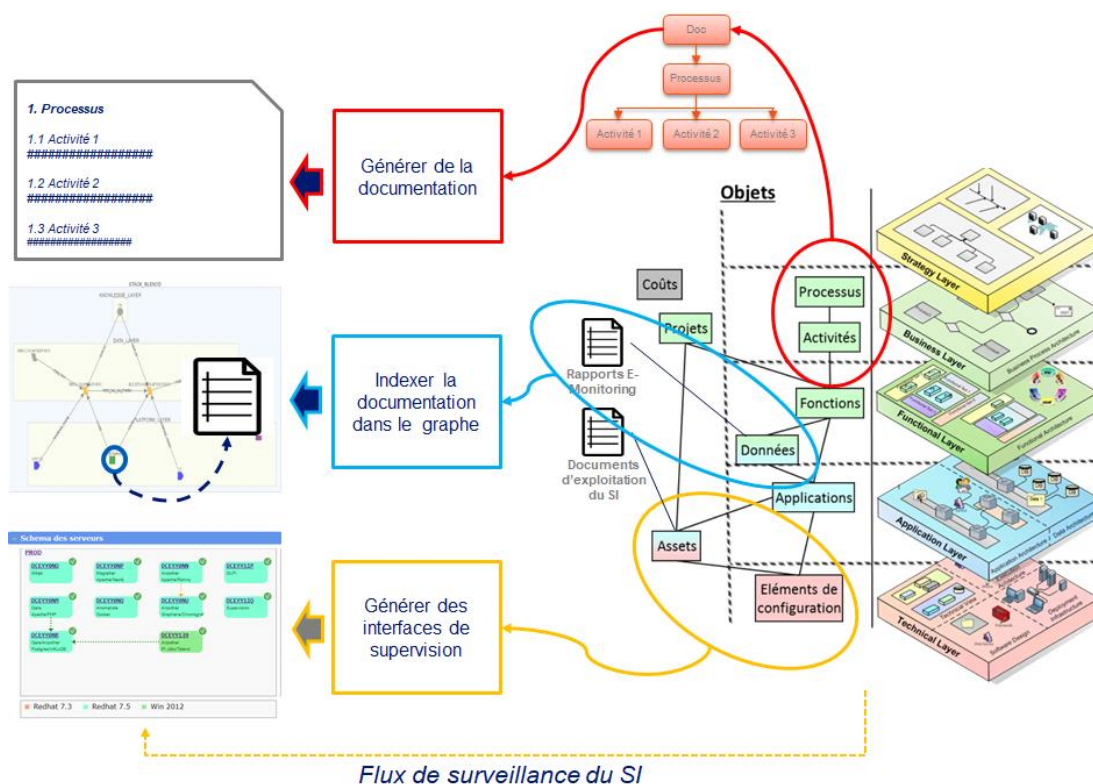


Figure 43 Les couches du SI et l'extension de MAGISTHER

## 6.1 ETENDRE LES CONCEPTS AU SEIN DU GRAPHE DE CONNAISSANCE

Avant de créer de nouveaux modules pour la supervision, nous devons enrichir le graphe existant des données de l'infrastructure réseaux. On s'inspire ici du Use Case de SFR qui présente une méthode de management des réseaux grâce à la cartographie des machines qui les composent. Avant la mise en place de ce Use Case, SFR explique que la mise en place de la maintenance de l'infrastructure prend plus d'une semaine à réaliser. Ce délai est lié à la nécessité de modéliser les impacts pour chacune des machines du réseau. Un réseau qui évolue vite parfois même quotidiennement.

Ils prévoient alors d'automatiser la modélisation de leur architecture réseau via la construction d'une cartographie. Cette cartographie s'appuie sur une base de graphes qui est une base de connaissances. Cette base de connaissances est alimentée par les données de plus de 30 machines. En représentant l'architecture réseaux sous forme de graphe, ils promettent la possibilité d'analyser et modéliser leurs réseaux pour un dépannage et une analyse d'impact rapide.

**SFR** Paris, France

Industry: Communications Use case: Network Management

- Second largest communications company in France
- Part of Vivendi Group, partnering with Vodafone

```

graph TD
    Service((Service)) -- DEPENDS_ON --> Router1((Router))
    Service -- DEPENDS_ON --> Router2((Router))
    Router1 -- DEPENDS_ON --> Switch1((Switch))
    Router1 -- DEPENDS_ON --> Switch2((Switch))
    Router2 -- DEPENDS_ON --> Switch1
    Router2 -- DEPENDS_ON --> Switch2
    Switch1 -- DEPENDS_ON --> FL1((Fiber Link))
    Switch1 -- DEPENDS_ON --> FL2((Fiber Link))
    Switch2 -- DEPENDS_ON --> FL3((Fiber Link))
    Switch2 -- DEPENDS_ON --> FL4((Fiber Link))
    FL4 -- DEPENDS_ON --> OC((Oceanfloor Cable))
  
```

- Infrastructure maintenance took one full week to plan, because of the need to model network impacts
- Needed rapid, automated "what if" analysis to ensure resilience during unplanned network outages identify weaknesses in the network to uncover the need for additional redundancy
- Network information spread across > 30 systems, with daily changes to network infrastructure Business needs sometimes changed very rapidly
- Flexible network inventory management system, to support modeling, aggregation & troubleshooting
- Single source of truth (Neo4j) representing the entire network
- Dynamic system loads data from 30+ systems, and allows new applications to access network data
- Modeling efforts greatly reduced because of the near 1:1 mapping between the real world and the graph
- Flexible schema highly adaptable to changing business requirements

Figure 44 Use case de SFR sur la gouvernance de l'information<sup>16</sup>

La modélisation des serveurs dans notre cartographie va permettre d'alimenter des modules déjà existants. La figure suivante décrit les concepts que nous avons ajoutés à notre base de connaissances. Notre cartographie a été enrichie des concepts : type d'OS (class OS), instance d'OS (class OS instance) et machine du réseau (class virtual machine).

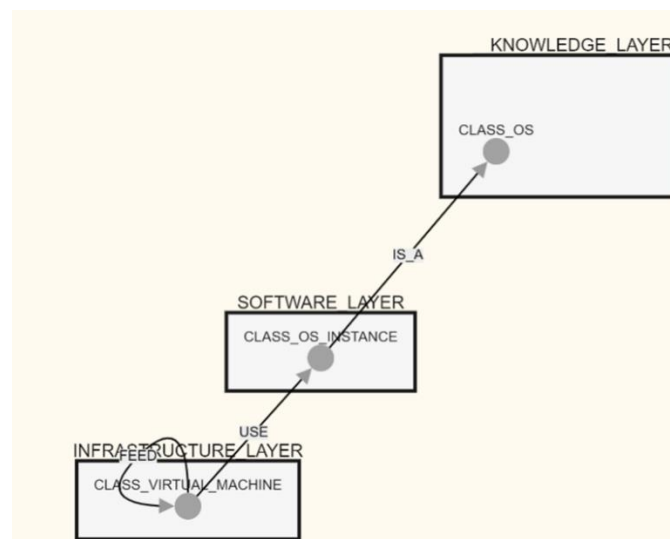


Figure 45 Les concepts ajoutés à la cartographie

<sup>16</sup> Source : <https://opencredo.com/case-studies/sfr/>



Aujourd'hui, chaque serveur est décrit dans un fichier. Ces fichiers décrivent les caractéristiques techniques des serveurs. Ses différentes caractéristiques sont éparpillées dans ce fichier XML et sont difficilement lisibles pour un utilisateur lambda.

Dans le module de catalogue, nous pourrons générer une liste complète des serveurs qui compose l'architecture du SI avec toutes les informations disponibles sur chacune des machines lisibles par l'humain.

The screenshot shows a web application interface for a server catalog. At the top, there is a blue header with a menu icon, the text 'Catalogues', and a user profile 'NUEVA.RENGA'. Below the header, the page title is 'INFORMATIQUE ET TELECOM/LES SERVEURS'. The main content is a table with the following columns: uid, stack, memory, os, vwsyst, ip, processor, description, company, and label. The table contains 38 rows of data, with the first 20 rows showing servers with 'MONITHER' stack and 'Microsoft Windows Se...' OS. The last 18 rows show servers with 'SODA-PROD' stack and 'Red Hat Enterprise Lin...' OS. At the bottom of the table, there is a status bar: 'Showing all 38 row(s)' on the left and 'Average: 0 Count: 0 Min: 0 Max: 0 Sum: 0 Data Filtered' on the right.

| uid                   | stack     | memory | os                        | vwsyst | ip                        | processor                 | description             | company                   | label       |
|-----------------------|-----------|--------|---------------------------|--------|---------------------------|---------------------------|-------------------------|---------------------------|-------------|
| /VIRTUAL_MACHINE/...  | MONITHER  | 4096   | Microsoft Windows Se...   | VMware | 56.48.98.72               | Intel(R) Xeon(R) CPU E... | PORCHEVILLE CPS-AF      | EDF                       | CPS-AF      |
| /VIRTUAL_MACHINE/...  | MONITHER  | 4096   | Microsoft Windows Se...   | VMware | 56.48.98.50               | Intel(R) Xeon(R) CPU E... | PORCHEVILLE CPS-PIPB    | EDF                       | CPS-PIPB    |
| /VIRTUAL_MACHINE/...  | MONITHER  | 8192   | Microsoft Windows Se...   | VMware | 56.48.98.15               | Intel(R) Xeon(R) CPU E... | CPS-PRISM               | CIH - DPHI - S2I Greno... | CPS-PRIS    |
| /VIRTUAL_MACHINE/...  | MONITHER  | 4096   | Microsoft Windows Se...   | VMware | 56.48.100.10              | Intel(R) Xeon(R) CPU E... | PORCHEVILLE SRV-AD1     |                           | SRV-AD      |
| /VIRTUAL_MACHINE/...  | MONITHER  | 4096   | Microsoft Windows Se...   | VMware | 56.48.100.11              | Intel(R) Xeon(R) CPU E... | PORCHEVILLE SRV-AD2     | EDF                       | SRV-AD      |
| /VIRTUAL_MACHINE/...  | MONITHER  | 4096   | Microsoft Windows Se...   | VMware | 56.48.4.10                | Intel(R) Xeon(R) CPU E... | PORCHEVILLE SRV-CSG     | EDF                       | SRV-CSG     |
| /VIRTUAL_MACHINE/...  | MONITHER  | 2048   | Microsoft Windows Se...   | VMware | 56.48.103.10              | Intel(R) Xeon(R) CPU E... | PORCHEVILLE SRV-FIL...  | EDF                       | SRV-FILESEF |
| /VIRTUAL_MACHINE/...  | MONITHER  | 2048   | Microsoft Windows Se...   | VMware | 56.48.6.10                | Intel(R) Xeon(R) CPU E... | PORCHEVILLE SRV-FTP     | EDF                       | SRV-FTP     |
| /VIRTUAL_MACHINE/...  | MONITHER  | 8192   | Microsoft Windows Se...   | VMware | 56.48.32.30               | Intel(R) Xeon(R) CPU E... |                         | EDF                       | SRV-JDB     |
| /VIRTUAL_MACHINE/...  | MONITHER  | 2048   | Microsoft Windows Se...   | VMware | 56.48.96.10               | Intel(R) Xeon(R) CPU E... | PORCHEVILLE SRV-LIC     | EDF                       | SRV-LIC     |
| /VIRTUAL_MACHINE/...  | MONITHER  | 4096   | Microsoft Windows Se...   | VMware | 56.48.32.40               | Intel(R) Xeon(R) CPU E... | PORCHEVILLE SRV-TO...   | EDF                       | SRV-TOO     |
| /VIRTUAL_MACHINE/...  | MONITHER  | 10240  | Microsoft Windows Se...   | VMware | 56.48.97.11               | Intel(R) Xeon(R) CPU E... | PORCHEVILLE SRV-VCT     |                           | SRV-VCT     |
| /VIRTUAL_MACHINE/...  | MONITHER  | 10240  | Microsoft Windows Se...   | VMware | 56.48.97.11               | Intel(R) Xeon(R) CPU E... | PORCHEVILLE SRV-VCT     |                           | SRV-VCT     |
| /VIRTUAL_MACHINE/...  | MONITHER  | 10240  | Microsoft Windows Se...   | VMware | 1.1.1.10                  | Intel(R) Xeon(R) CPU E... | PORCHEVILLE SRV-VE...   | EDF                       | SRV-VEEA    |
| /VIRTUAL_MACHINE/...  | MONITHER  | 4096   | Microsoft Windows Se...   | VMware | 56.48.7.10                | Intel(R) Xeon(R) CPU E... |                         | EDF                       | SRV-WEB-AI  |
| /VIRTUAL_MACHINE/...  | MONITHER  | 8192   | Microsoft Windows Se...   | VMware | 56.48.68.30               | Intel(R) Xeon(R) CPU E... |                         | EDF                       | SRV-WEB-PI  |
| /VIRTUAL_MACHINE/...  | MONITHER  | 4096   | Microsoft Windows Se...   | VMware | 56.48.101.10              | Intel(R) Xeon(R) CPU E... | PORCHEVILLE SRV-WS...   | EDF                       | SRV-WSL     |
| /VIRTUAL_MACHINE/S... | SODA-PROD | 15876  | Red Hat Enterprise Lin... | VMware | 10.129.10.57/10.129.1...  | Intel(R) Xeon(R) CPU E... | x86_64/00-00-09 04:5... |                           | dceyy0np    |
| /VIRTUAL_MACHINE/S... | SODA-PROD | 7812   | Red Hat Enterprise Lin... | VMware | 10.129.10.58/10.129.1...  | Intel(R) Xeon(R) CPU E... | x86_64/00-00-09 04:5... |                           | dceyy0np    |
| /VIRTUAL_MACHINE/S... | SODA-PROD | 7812   | Red Hat Enterprise Lin... | VMware | 10.129.10.59/10.129.1...  | Intel(R) Xeon(R) CPU E... | x86_64/00-00-09 04:2... |                           | dceyy0np    |
| /VIRTUAL_MACHINE/S... | SODA-PROD | 7812   | Red Hat Enterprise Lin... | VMware | 10.129.10.60/10.129.1...  | Intel(R) Xeon(R) CPU E... | x86_64/00-00-09 04:5... |                           | dceyy0np    |
| /VIRTUAL_MACHINE/S... | SODA-PROD | 7812   | Red Hat Enterprise Lin... | VMware | 172.18.0.1/172.17.0.1/... | Intel(R) Xeon(R) CPU E... | x86_64/00-00-09 04:5... |                           | dceyy0np    |
| /VIRTUAL_MACHINE/S... | SODA-PROD | 15876  | Red Hat Enterprise Lin... | VMware | 10.129.10.62/10.129.1...  | Intel(R) Xeon(R) CPU E... | x86_64/00-00-09 04:5... |                           | dceyy0np    |
| /VIRTUAL_MACHINE/S... | SODA-PROD | 7812   | Red Hat Enterprise Lin... | VMware | 10.129.10.63/10.129.1...  | Intel(R) Xeon(R) CPU E... | x86_64/00-00-09 04:5... |                           | dceyy0np    |
| /VIRTUAL_MACHINE/S... | SODA-PROD | 16384  | Microsoft Windows Se...   | VMware | 10.129.124.16             | Intel(R) Xeon(R) CPU E... |                         | EDF - CSPITO              | DCEYY13     |
| /VIRTUAL_MACHINE/S... | SODA-PROD | 7812   | Red Hat Enterprise Lin... | VMware | 10.129.10.157/10.129...   | Intel(R) Xeon(R) CPU E... | x86_64/00-00-09 04:5... |                           | dceyy1if    |



© QARE 2020

Figure 46 Catalogue des données des serveurs

Le module de cartographie va permettre de tracer le lignage entre les différents systèmes. L'illustration suivante nous permet d'affirmer qu'il existe un serveur (dceyy0np en surbrillance) qui utilise une installation RedHat Linux 7.7. Cette installation est elle-même une instance du système RedHat Linux décrite dans notre couche de connaissances (Knowledge Layer). En ciblant notre recherche sur ce dernier élément, nous sommes même capables de lister l'ensemble des serveurs qui utilisent une instance de cette OS.

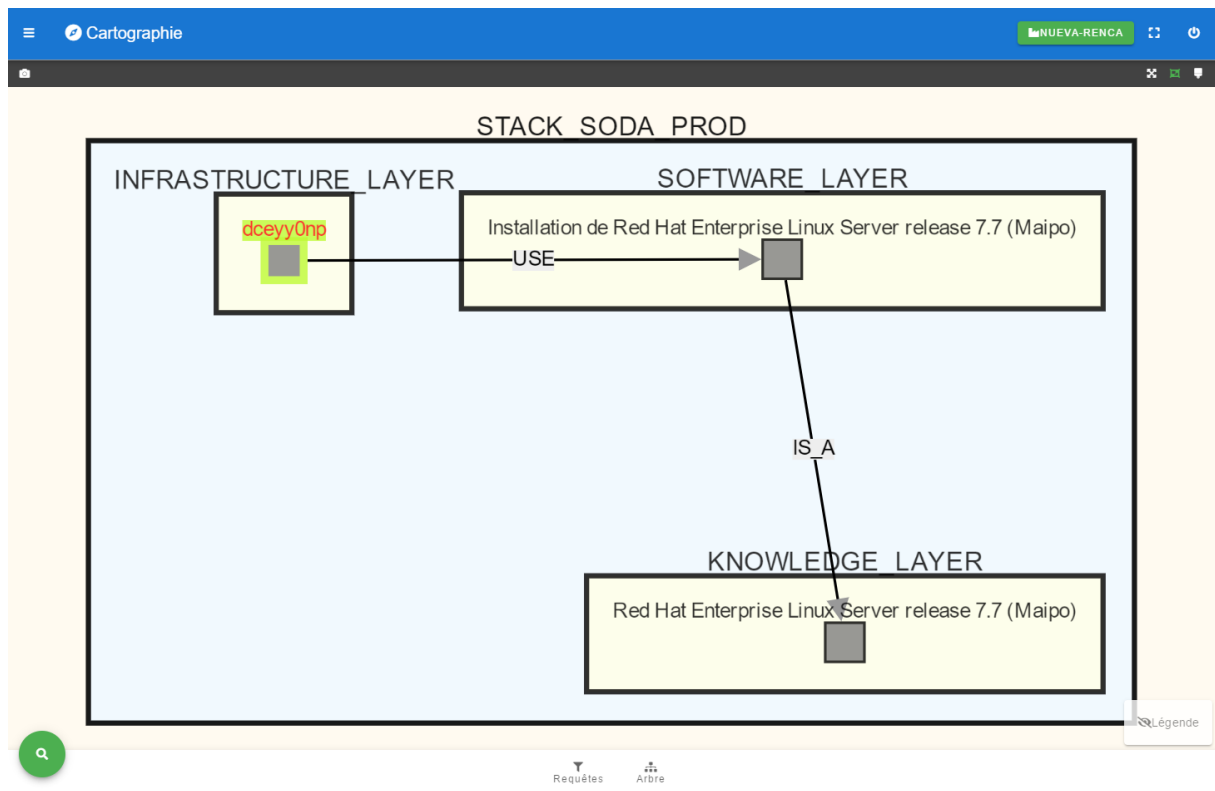


Figure 47 Implantation des données des serveurs dans le module cartographie existant

## 6.2 GENERER DES INTERFACES DE SUPERVISION

### 6.2.1 Contexte

Dans notre unité, nous exploitons des données provenant de plusieurs machines (serveurs, historian...). Sur ces machines, plusieurs scripts ou procédures se déclenchent de façon automatique à des intervalles qui diffèrent selon leurs fonctions.

Avec un nombre de procédures croissant sur des machines différentes, il est de plus en plus difficile de s'assurer du bon fonctionnement de l'ensemble du système.

Il nous est également difficile et fastidieux de vérifier une par une toutes les machines pour connaître leur état.

Nous avons donc mis en place, des écrans de supervisions qui s'appuient sur des données exportées de chaque machine utilisée pour afficher en temps réel : l'état de machine, l'historique des processus par machine ...

Cet écran se présente sous cette forme :

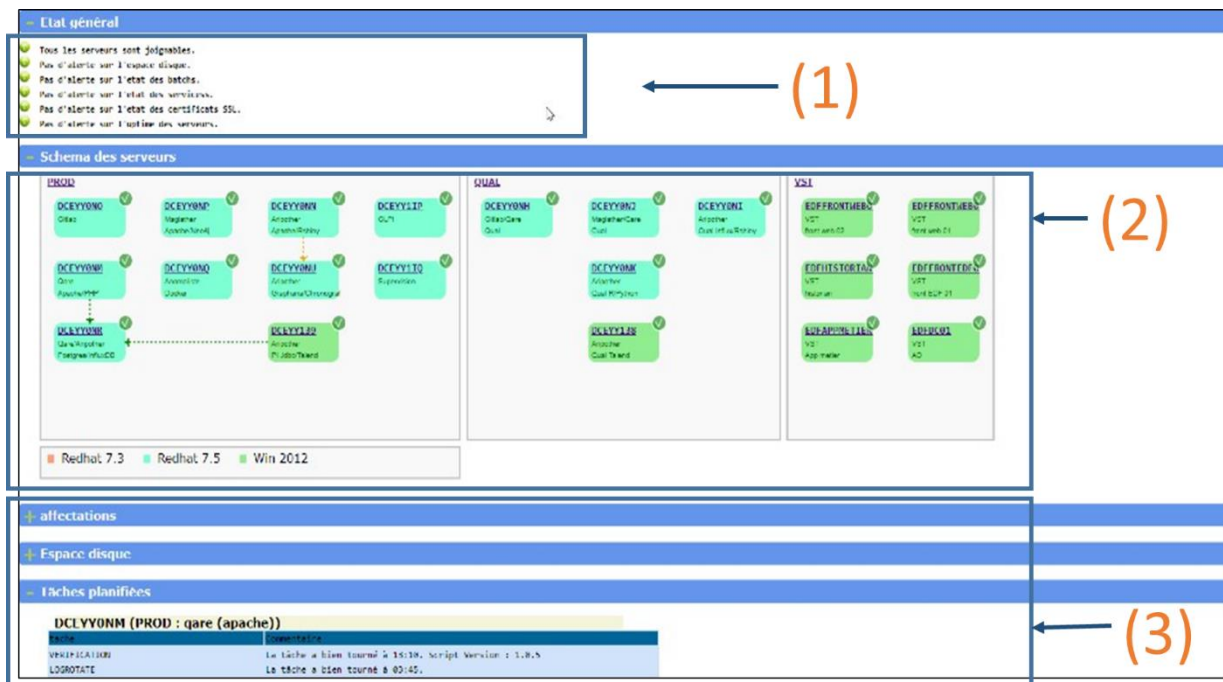


Figure 48 Ecran de supervision EDF Ingeum

Cet écran représente l'état général des serveurs (1) ainsi qu'une représentation schématique des machines composant les réseaux d'INGEUM (2). Dans ce schéma, nous pouvons voir les relations entre les machines, leur état ainsi que des groupements par infrastructure. Cet écran inclut des onglets répertoriant des informations complémentaires sur chacun des serveurs représentés (3).

### 6.2.2 Problématique

Cet écran a été réalisé et est maintenu manuellement par un de nos prestataires. A chaque modification du SI (Système d'information), nous sommes obligés d'adapter toute la structure de l'écran afin de qu'il prenne en compte les changements.

### 6.2.3 Solution

Cette partie du projet vise donc à intégrer la couche technique de la modélisation du SI dans notre graphe afin de générer des interfaces de supervisions automatiquement.

Ces interfaces seront alimentées par les données sur les machines de la même façon qu'est alimenté l'écran actuel.

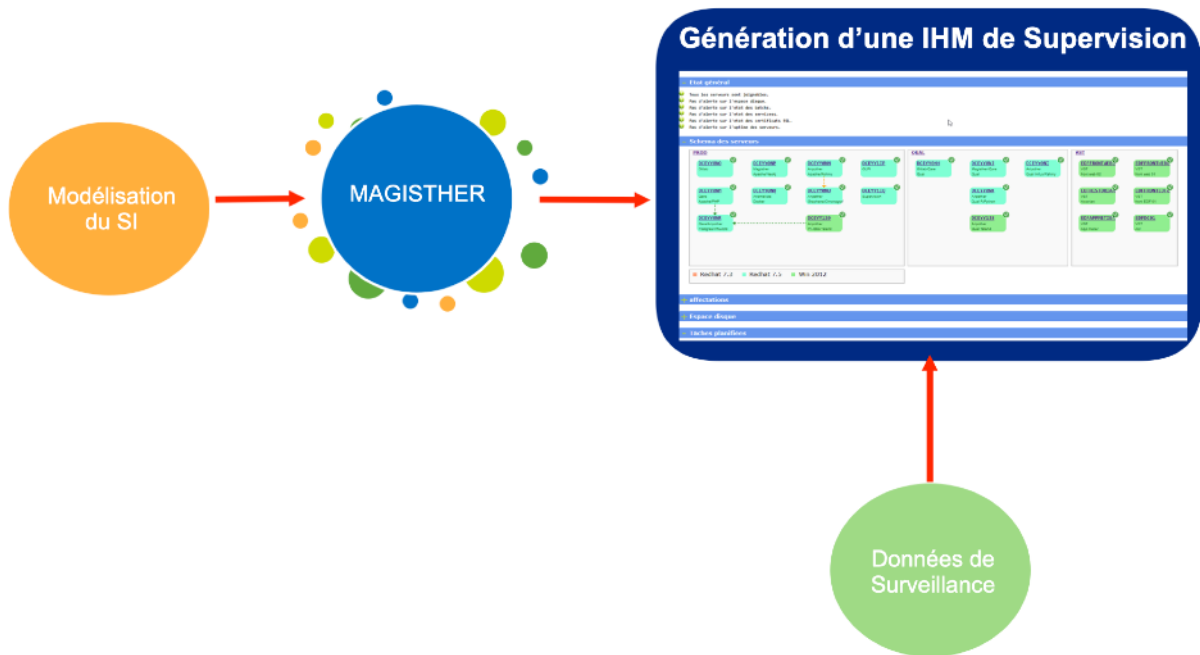


Figure 49 Schématisation de la génération des écrans de supervision

L'avantage premier de cette méthode est d'avoir des interfaces de supervision qui s'adaptent au changement dans la modélisation du SI.

Nous incorporons au module trois manières d'afficher les données :

- Une méthode topologique,
- Une visualisation plus proche de l'écran initialement conçu par notre prestataire (mode carte),
- Une représentation sous forme de tableau.

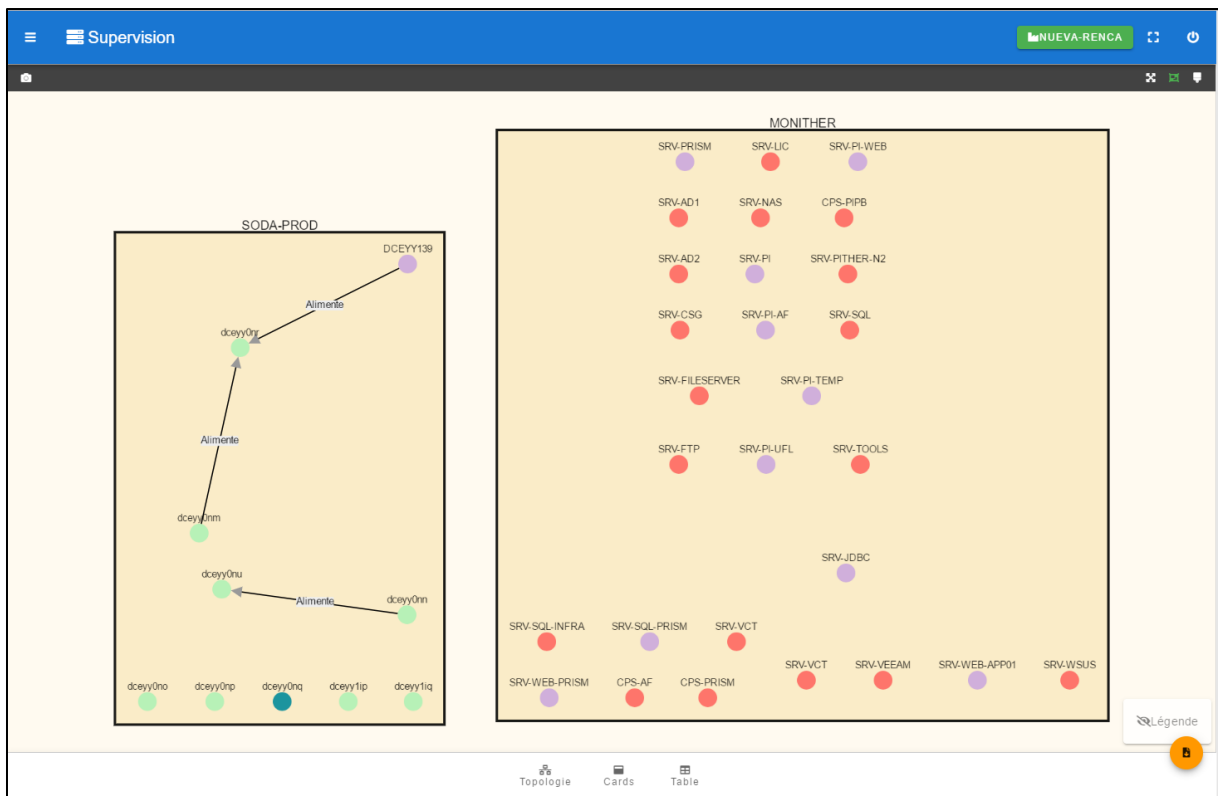


Figure 50 Représentation graphique des serveurs

Dans ce mode topologique, chaque nœud représente un serveur de l'architecture d'Ingeum présent dans notre base de connaissances. Nous pouvons voir les différentes connexions entre les machines ou encore à quel environnement elles appartiennent grâce au cluster (groupement jaune sur la capture).

En cliquant sur un nœud, nous pouvons accéder à toutes les métadonnées rattachées au serveur sélectionné.

| Types  | Valeurs                                     |
|--------|---|
| id     | SRV-FTP-2019-12-08-09-58-03                 |
| label  | SRV-FTP                                     |
| osName | Microsoft Windows Server 2008 R2 Enterprise |
| color  | #FF6961                                     |
| parent | MONITHER                                    |

Figure 51 Panel d'information sur un serveur

Le mode carte est une représentation synthétique sous forme de carte où chacune d'entre elles représente un serveur.

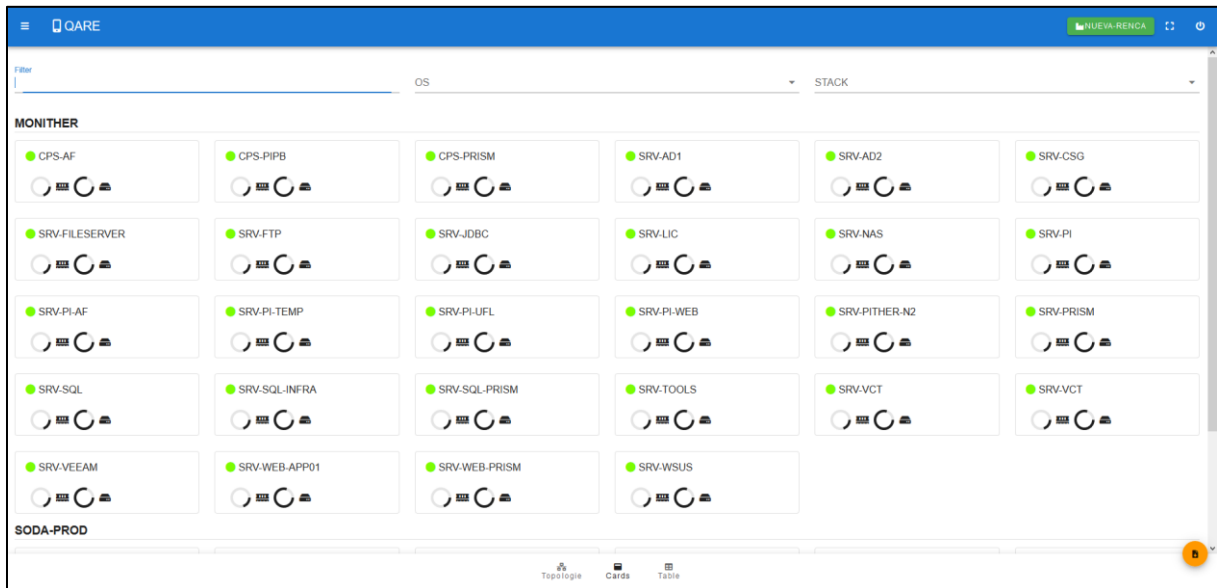


Figure 52 Représentation synthétique des serveurs du SI

Nous avons incorporé un filtrage des serveurs grâce à leurs attributs (OS, environnement...). Ainsi qu'une option de filtrage par nom comme présenté sur la figure suivante.

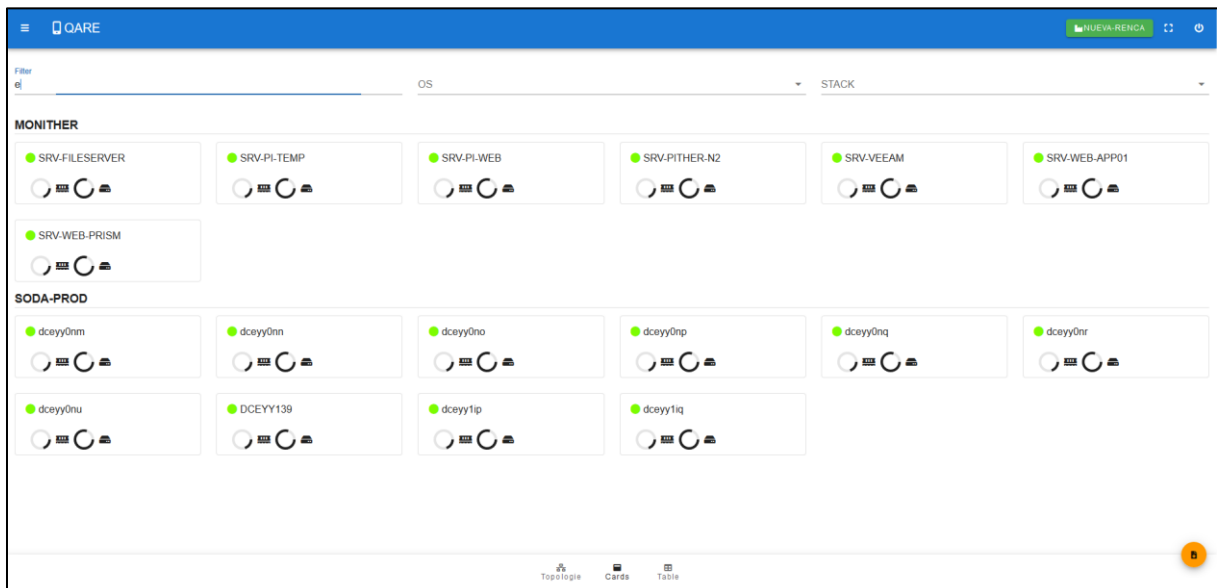


Figure 53 Représentation synthétique des serveurs du SI (Filtrage)

Une version tabulaire représentant toutes les informations concernant les serveurs de notre base de graphes avec une recherche par nom afin d'accéder rapidement aux données voulues.

| Label                     | group | Identifiant                        | group | OS  | group |
|---------------------------|-------|------------------------------------|-------|---|-------|
| — data.parent: MONITHER X |       |                                    |       |   |       |
| SRV-SQL-INFRA             |       | SRV-SQL-INFRA-2019-12-03-05-00-06  |       | Microsoft Windows Server 2008 R2 Enterprise |       |
| CPS-AF                    |       | CPS-AF-2020-03-05-05-00-07         |       | Microsoft Windows Server 2008 R2 Enterprise |       |
| CPS-PRISM                 |       | CPS-PRISM-2019-12-03-05-00-15      |       | Microsoft Windows Server 2008 R2 Enterprise |       |
| SRV-AD1                   |       | SRV-AD1-2019-12-03-05-00-05        |       | Microsoft Windows Server 2008 R2 Enterprise |       |
| SRV-AD2                   |       | SRV-AD2-2020-02-19-05-00-06        |       | Microsoft Windows Server 2008 R2 Enterprise |       |
| SRV-CSG                   |       | SRV-CSG-2019-12-03-05-00-12        |       | Microsoft Windows Server 2008 R2 Enterprise |       |
| SRV-FILESERVER            |       | SRV-FILESERVER-2019-12-03-05-00-12 |       | Microsoft Windows Server 2008 R2 Enterprise |       |
| SRV-FTP                   |       | SRV-FTP-2019-12-06-09-58-03        |       | Microsoft Windows Server 2008 R2 Enterprise |       |
| SRV-JDBC                  |       | SRV-JDBC-2019-12-03-05-00-10       |       | Microsoft Windows Server 2012 R2 Standard   |       |
| SRV-LIC                   |       | SRV-LIC-2019-12-03-05-00-13        |       | Microsoft Windows Server 2008 R2 Enterprise |       |
| SRV-NAS                   |       | SRV-NAS-2019-12-03-05-00-11        |       | Microsoft Windows Server 2008 R2 Enterprise |       |
| SRV-PI                    |       | SRV-PI-2019-12-03-05-00-09         |       | Microsoft Windows Server 2012 R2 Standard   |       |
| SRV-PI-AF                 |       | SRV-PI-AF-2019-12-03-05-00-13      |       | Microsoft Windows Server 2012 R2 Standard   |       |
| SRV-PI-TEMP               |       | SRV-PI-TEMP-2019-12-03-05-00-04    |       | Microsoft Windows Server 2012 R2 Standard   |       |
| SRV-PI-UFL                |       | SRV-PI-UFL-2019-12-03-05-00-12     |       | Microsoft Windows Server 2012 R2 Standard   |       |

Figure 54 Représentation des serveurs sous forme de tableau téléchargeable dans au format Excel

Les différents modes sont également compatibles avec une utilisation sur smartphone ou tablette. Ceci permet d'envisager de consulter l'état des serveurs depuis son mobile lors d'un déplacement à l'étranger par exemple.

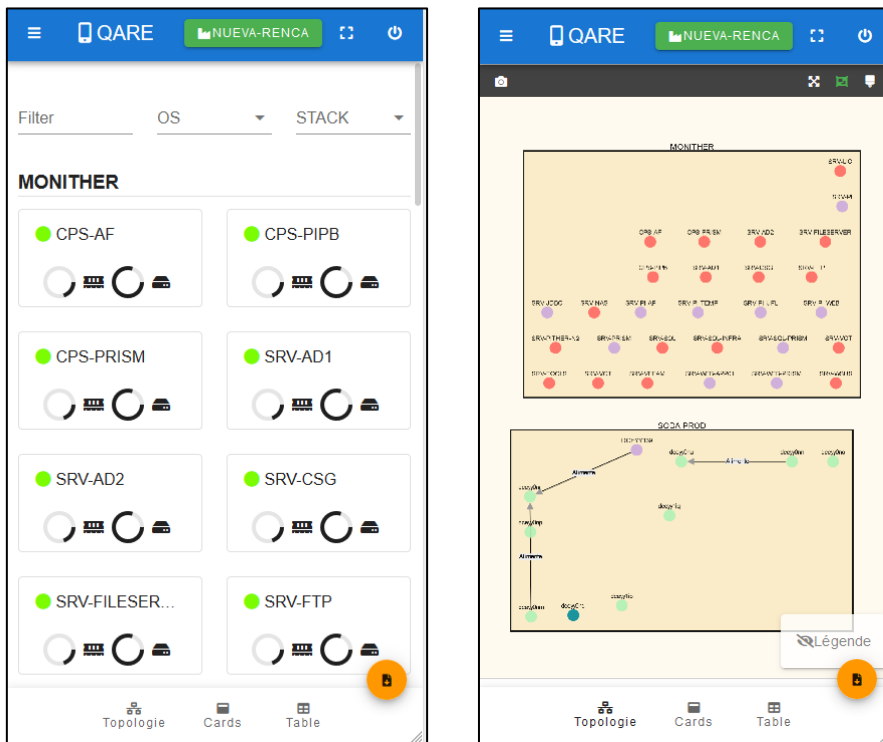


Figure 55 Module de supervision version mobile

## 6.3 INDEXER LA DOCUMENTATION

### 6.3.1 Contexte

EDF, comme bon nombre d'entreprises, met en œuvre de nombreux process métiers s'appuyant sur la documentation.

Elle est exploitée dans le cadre de :

- La vie des équipes : Compte-rendu de réunion, planning, communication, recrutement, mutation ...);
- La gestion de projet : (schéma, cahier des charges, organigramme...);
- L'architecture du SI (configuration réseaux, matrice de flux...);
- La maintenance (planning d'astreinte, process en cours de traitement...).

Et comme bon nombre d'autres structures, EDF est dans une base où la documentation s'entasse, vieillit, s'éparpille nécessitant une mise à jour et un référencement pour les créateurs de ces documents.

### 6.3.2 Problématique

INGEUM et la DTEAM n'est pas exclu du contexte documentaire présenté précédemment.

Accéder et trouver un document relatif à une application ou un serveur prend du temps. En effet, les sources documentaires éparpillées ne simplifient pas la prise en main de la documentation ni la recherche d'informations sur un sujet donné.

### 6.3.3 Solution

Pour répondre à cette problématique, nous nous appuyons sur une étude réalisée par IBM/EDF sur « la génération composite de document à partir de sources documentaires de l'entreprise » illustrée par le schéma suivant :



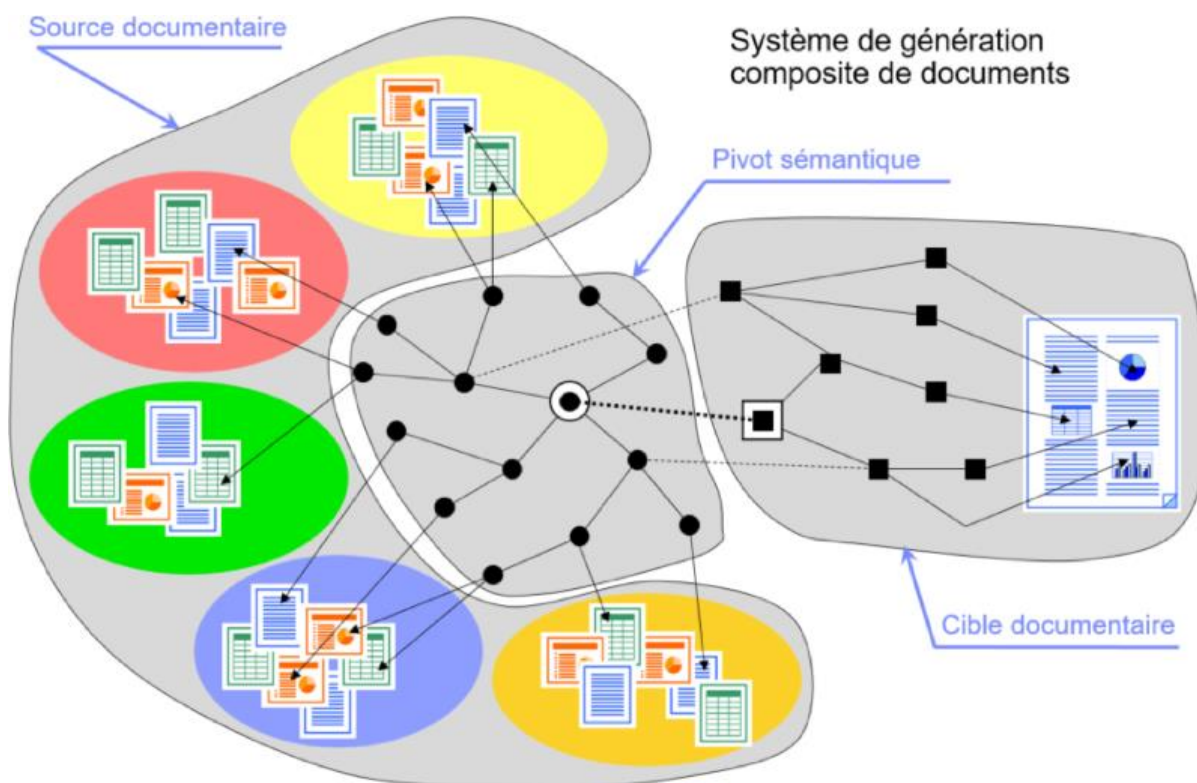


Figure 56 Système de génération composite de document par Samuel PARFOURU et Pascal NEGROS

La déclinaison de cette approche dans le contexte de MAGISTHER peut être illustrée dans la figure suivante :

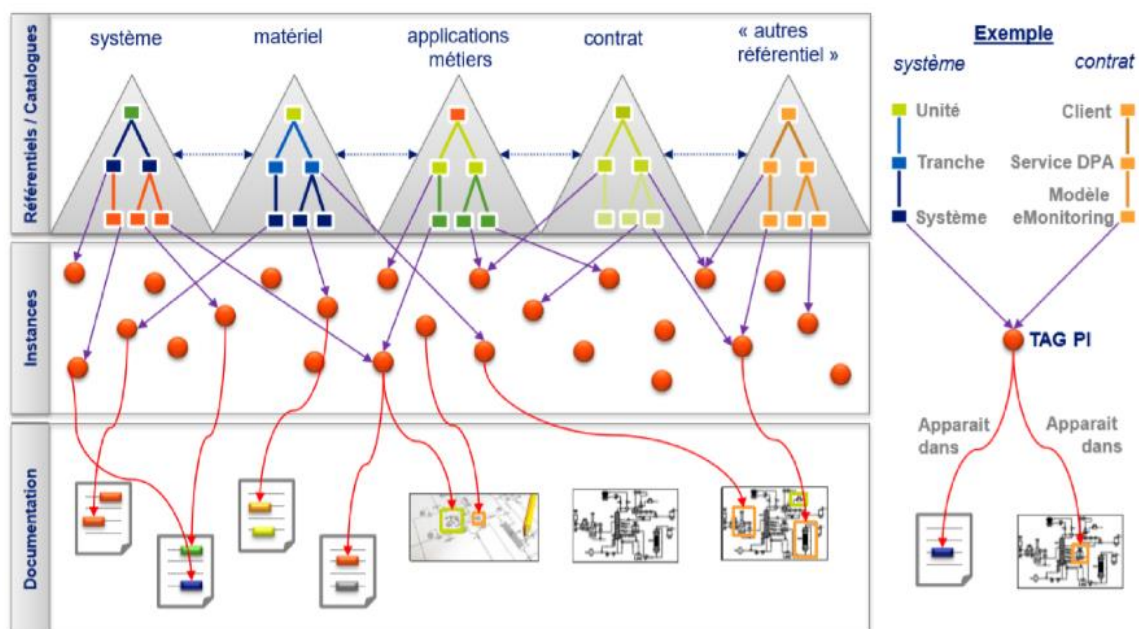


Figure 57 Cartographie du SI dans MAGISTHER

Ces deux documents nous soutiennent que nous pouvons ajouter des documents en tant que nœuds dans notre cartographie et les lier à la modélisation de notre SI.

Nous pouvons alors rechercher un élément dans notre cartographie des données et en un simple « click » accéder à l'ensemble de la documentation qui mentionne la donnée recherchée. Cette méthode permettra de fortement alléger le travail de recherche documentaire puisque toutes les informations seront cartographiées et accessibles via la même interface quel que soit le type de document (fichier de log, plan 2D, plan 3D...).

Nous avons donc incorporé dans notre maquette un visualiseur de document. A l'aide d'un navigateur sous forme d'arbre de connaissances (1), nous pouvons cibler un concept et afficher dans notre navigateur, les différents documents qui décrivent ce dernier (2), mélangeant ainsi schéma de conception, document d'exploitation...

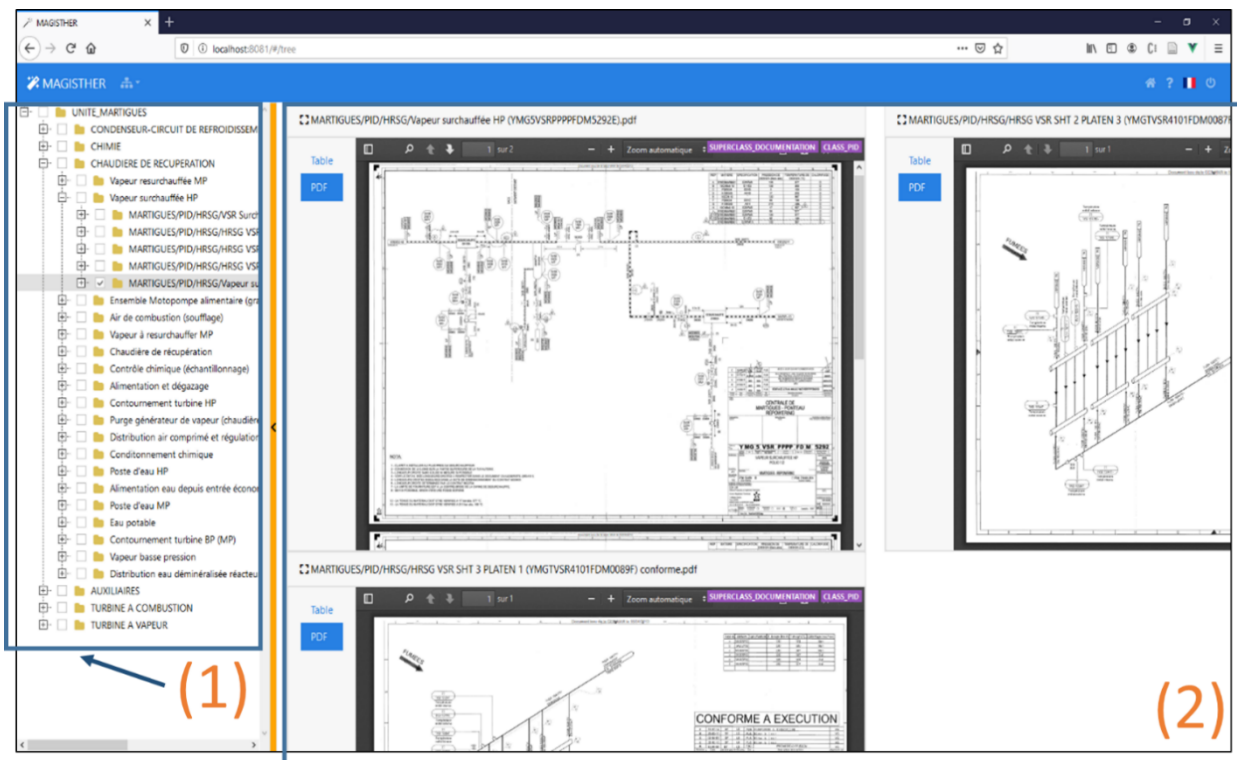


Figure 58IHM de visualisation de document

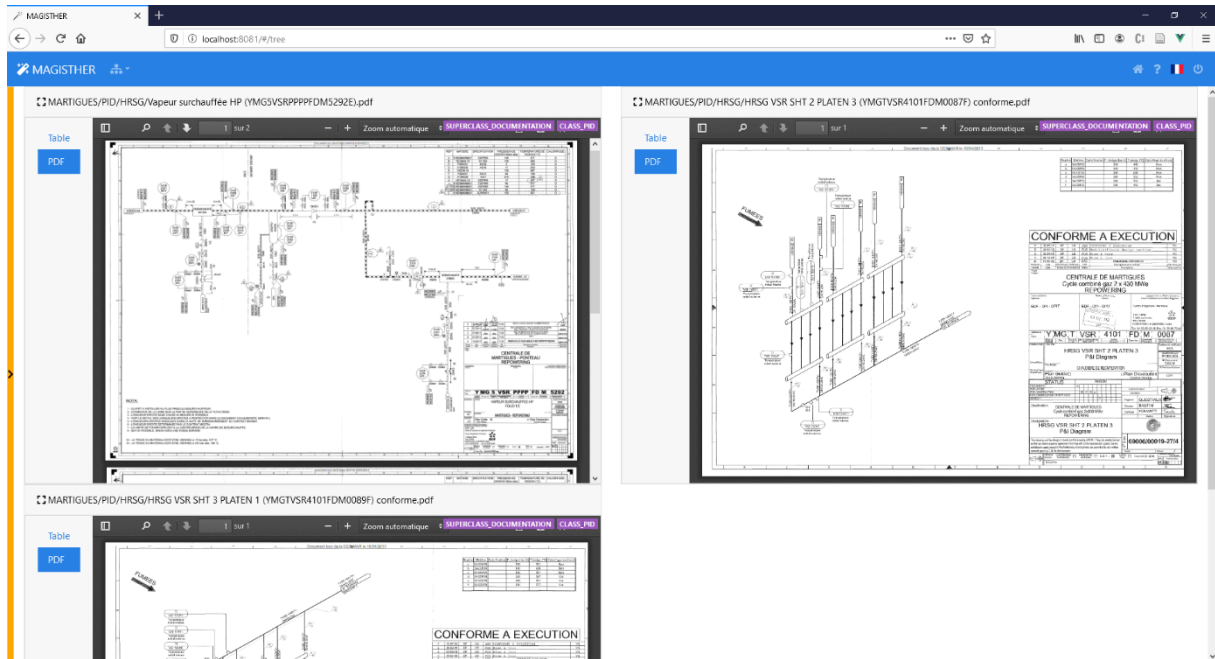


Figure 59 IHM de visualisation de document centré sur la documentation

## 6.4 GENERER DE LA DOCUMENTATION

### 6.4.1 Contexte

Comme tous les SI, celui d'INGEUM est décrit par des documents. Dans ces documents, on peut trouver l'ensemble des caractéristiques de notre architecture comme la configuration des firewalls ou la matrice des flux entre les serveurs. Certains métiers exploitent cette documentation ce qui implique qu'elle doit être maintenue à jour.

### 6.4.2 Problématique

La transformation numérique que vit INGEUM, entraîne une évolution rapide et régulière de son SI et donc une péremption plus rapide de la documentation qui le décrit. Maintenir à jour ces documents devient une tâche de plus en plus ardue et consommatrice de temps.

### 6.4.3 Solution

Notre objectif ici est de concevoir un système de création « intelligente » et automatique de document, à partir des informations de la base de graphes.

En parcourant le graphe et en ajoutant un modèle de mise forme PDF, nous serons capables d'exploiter les relations entre les éléments du référentiel pour construire une documentation métier automatiquement.

Dans l'illustration ci-dessous, nous avons schématisé le process de génération de documents à partir d'une configuration réseau.

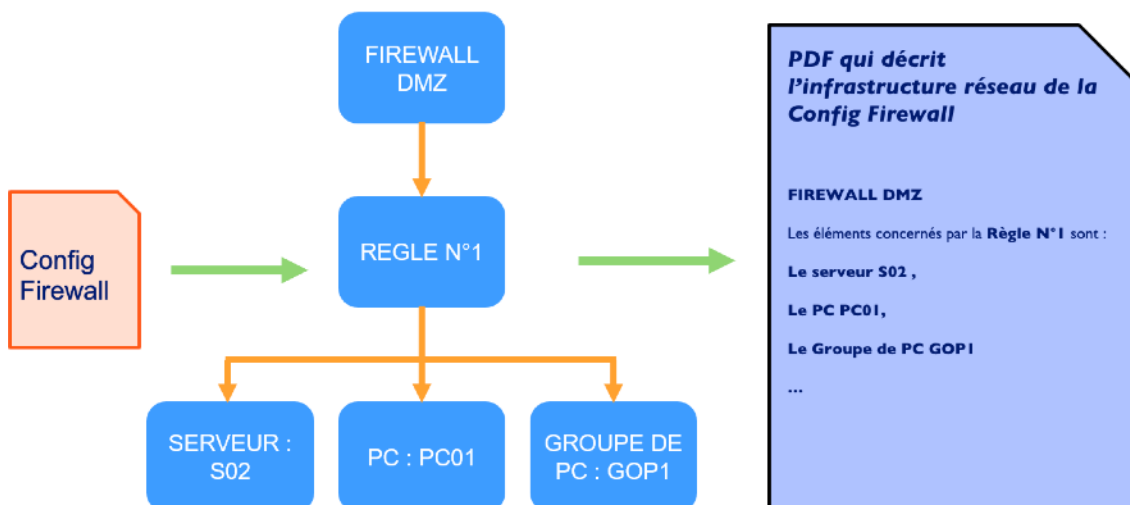


Figure 60 Principe de Génération automatique de document

Ici, nous intégrons une configuration réseau dans notre base de graphes.

Une fois les données structurées dans le graphe, nous pourrions exporter les informations dans un document PDF explicitant la configuration réseau élément par élément.

Cette méthode pourra être utilisée pour produire des documents pour le métier ainsi que pour le maintien du SI par exemple ; et même, permettre de généraliser et automatiser une multitude de création de documentation au sein de notre unité.

Dans le nouveau module de supervision, nous avons incorporé une génération d'un document PDF lisible par l'humain. Ce document incorpore l'ensemble des données concernant l'architecture réseau d'INGEUM contenu dans notre base de connaissance.

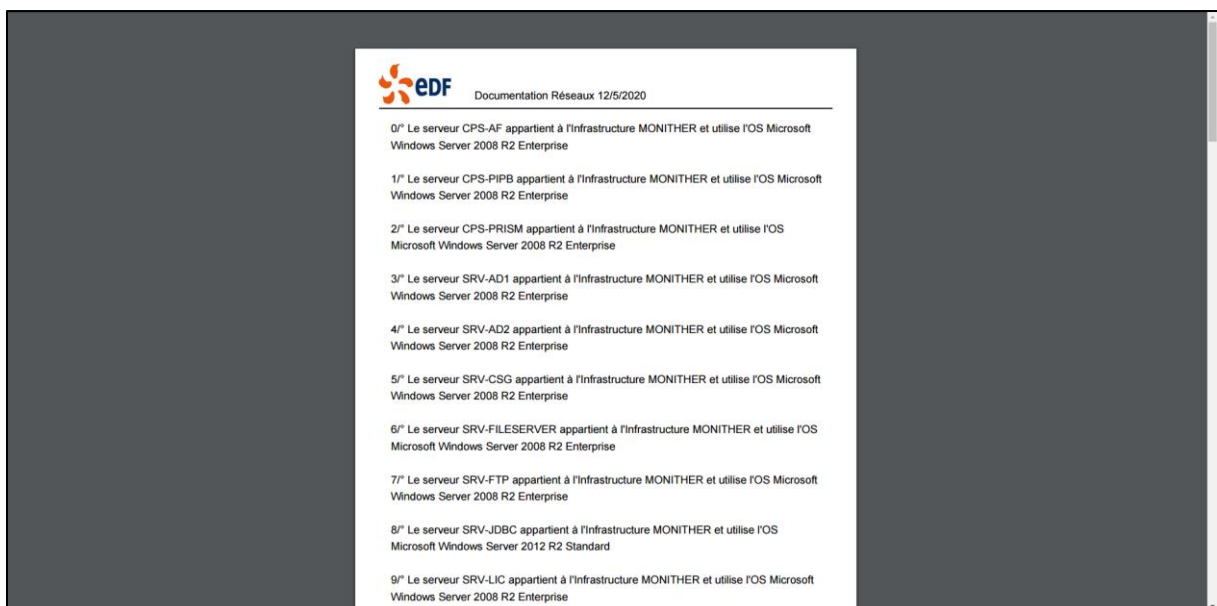


Figure 61 (Génération d'un document basé sur les éléments du graphe)

Nous avons incorporé ce générateur uniquement dans le module de supervision. Mais ce développement prouve qu'il est possible de générer de la documentation en fonction des données de notre cartographie. Il est donc envisageable et réalisable d'incorporer cette génération pour chaque concept de la cartographie et ainsi maintenir une documentation automatique à l'échelle de l'équipe 2i mais aussi d'INGEUM.

## 7 DISCUSSION & PERSPECTIVES

---

### 7.1 BILAN

Mes trois années d'alternance ont été rythmées par la démarche MAGISTHER.

Nous avons réussi en 3 ans à mettre en œuvre une démarche changeant radicalement la manière dont les métiers exploitent l'information. Nous recevons de plus en plus de demandes nous invitant à organiser des ateliers qui présentent notre portail et comment s'en servir au quotidien.

Certains prestataires ont déjà été formés et utilisent l'outil de cartographie pour connaître le lignage des TAGs concernés par une demande de modification.

Les bases de graphes, les technologies d'ETL, le développement web et la gouvernance de données étaient des sujets peu maîtrisés ou inconnus pour moi. Ces trois années d'apprentissage m'ont permis de devenir autonome sur l'ensemble de ces sujets et de prendre en charge l'extension de MAGISTHER.

Nous avons établi une maquette fonctionnelle démontrant la faisabilité des extensions de la démarche MAGISTHER appliquée au cas concret de l'architecture réseaux. Une phase de démonstration et de discussions avec les utilisateurs finaux est déjà en cours afin d'établir les fonctionnalités attendues des modules pour en produire des versions industrielles en production.

### 7.2 PROCHAINES ETAPES

La démarche MAGISTHER est une démarche en évolution constante. Ainsi, d'autres pistes d'exploitation peuvent être envisagées.

#### 7.2.1 Vérification des serveurs automatisée par le graphe

Sur chaque serveur de notre architecture réseau, tournent tous les jours, des scripts de vérification. Ces scripts lancent une analyse complète des tâches planifiées sur les serveurs ainsi que l'état de ces derniers. La résultante de cette analyse est générée sous forme de fichier (XML ou texte).

Ce sont ces fichiers que nous exploitons pour annexer la documentation aux nœuds du graphe dans le module de supervision.

A l'heure actuelle, ces scripts de vérification sont écrits à la main par l'un de nos prestataires et implantés sur chaque serveur. Chaque script contient les mêmes commandes avec pour seules variantes les noms des serveurs et les tâches à surveiller.

Les noms des serveurs sont des informations déjà contenues dans notre cartographie. Nous pourrions ajouter les tâches à surveiller en tant que nœuds dans notre base de connaissances et créer une relation entre chaque tâche et le serveur qui exploite cette dernière.

Nous avons déjà conçu un générateur de document prenant comme source les informations du graphe. En adaptant le type de fichier de sortie ainsi que le contenu du fichier, nous serons capables de créer automatiquement un fichier de vérification par serveur en fonction des tâches qu'ils lui sont associées dans le graphe.

En incorporant ces nouvelles informations dans notre base de connaissances, nous serons capables de générer des scripts de vérification pour chaque serveur et ainsi rendre la vérification des serveurs automatique et commandée par le graphe.

## 7.2.2 Supervision des scripts ETL

Au cours de ce rapport, nous avons évoqué de nombreux scripts permettant la mise en valeur des données, la construction du graphe. Aujourd'hui, énormément de scripts tournent de façon régulière et permettent à MAGISTHER de fonctionner et d'alimenter tous les modules cités précédemment. Ils s'exécutent indépendamment les uns des autres à des rythmes et horaires différents.

A l'heure actuelle, seul le responsable de l'infrastructure est capable de dire si un des scripts qui tourne sur ce dernier s'est exécuté avec succès ou pas. Pour ce faire, il doit vérifier un par un les logs de chaque script. Il a donc une charge de travail lié à la surveillance de ce script afin de détecter et résoudre les problèmes le plus rapidement possible.

Pour simplifier cette vérification, nous avons évoqué l'idée de créer un module de surveillance des scripts au sein de MAGISTHER. Ce module permettra de visualiser rapidement l'intégralité des scripts ainsi que leur état (échouer, exécuter sans problème, en cours d'exécution).

Nous exploiterons notre forge logicielle Gitlab afin de versionner et conserver les différents scripts dans un répertoire dédié. Nous incorporerons un Job qui viendra exploiter les métadonnées de l'ensemble de ce job afin d'afficher dans le module les différentes informations

sur les scripts (étapes principales, objectifs, fichiers sources, fichiers de sortie, emplacement...).

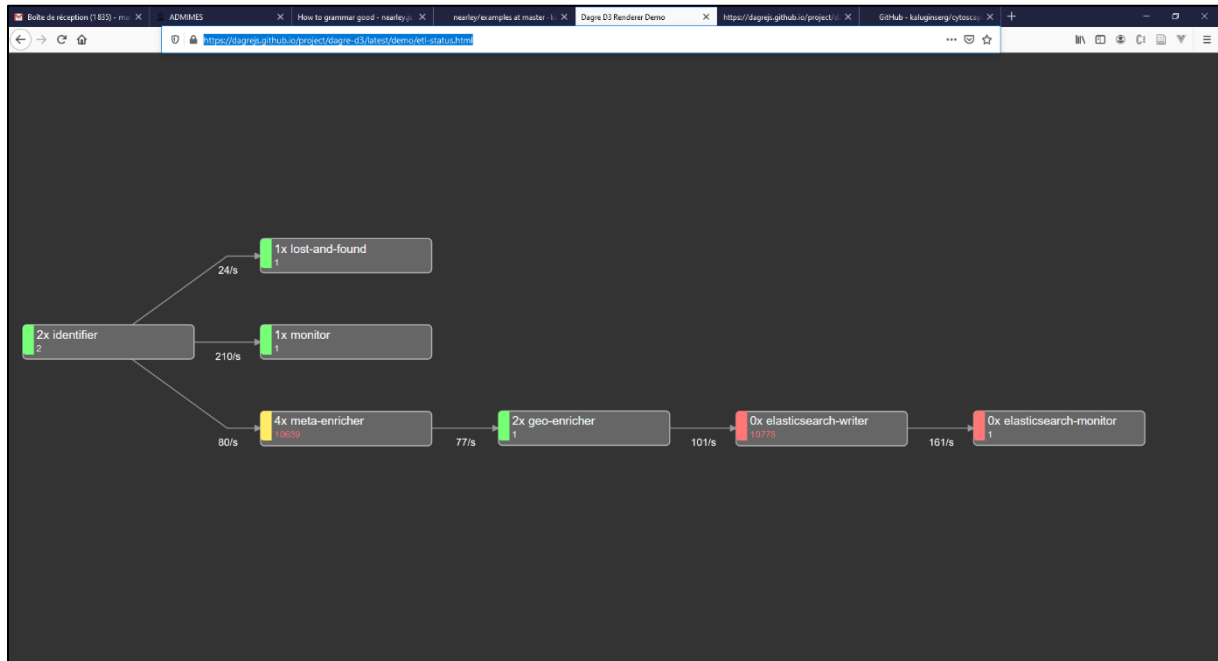


Figure 62 IHM exploitant DagreJS<sup>17</sup> répondant au besoin de supervision d'un script

Un moteur capable de répondre à cet objectif est réalisable et un prototype de ce dernier a déjà été testé sur des scripts de test.

Il n'y a ici aucun verrou technique afin d'ajouter ce module. Nous avons déjà initié la démarche de regroupement des différents scripts et d'y ajouter une tâche planifiée permettant de surveiller les différents scripts. Cette tâche planifiée est basée sur le moteur cité plus haut dont une maquette en cours d'exploitation.

Je vais prendre en charge cette intégration dans les mois qui viennent. Je vais travailler avec plusieurs prestataires afin d'organiser la mise en place du nouveau répertoire et de l'organisation des scripts à l'intérieur de ce dernier.

### 7.2.3 Analyse d'impact

Nous avons construit une cartographie des données de notre SI et ainsi nous avons pu reconstituer la chaîne d'approvisionnement des données.

Aujourd'hui, lorsqu'un site doit supprimer ou modifier un TAG (le renommer par exemple), il nous envoie une demande afin que nous puissions vérifier que le TAG n'est pas impliqué dans un process actif (que le TAG n'alimente pas un autre TAG lié à une application active). Avec le

<sup>17</sup> Source : <https://dagrejs.github.io/project/dagre-d3/latest/demo/etl-status.html>



module de cartographie que nous avons créé, nous pouvons rechercher le TAG à supprimer, trouver son lignage et ainsi déduire son implication dans une chaîne d’approvisionnement active ou non.

Dans le cas où le site nous transmet une liste de plusieurs TAGs, nous devons rechercher le lignage dans la cartographie TAG par TAG jusqu’à trouver les connexions de toute la liste fournie. Le module de cartographie n’est pas adapté pour l’analyse d’impact sur un ensemble de TAGs mais intègre toutes les données nécessaires afin de générer un module dont le seul objectif serait de faciliter l’analyse d’impact.

Nous avons donc conçu la maquette ci-dessous. Nous avons en (1) un filtre de recherche permettant de spécifier le type de données recherché ainsi que l’infrastructure contenant nos TAGs cibles. Afin de supprimer la redondance dans l’analyse des TAGs, nous avons ajouté un champ permettant de copier-coller la liste des TAGs à analyser directement depuis la demande des sites (2). Ces informations vont générer automatiquement une requête sur notre base de graphes ayant pour résultante l’ensemble des éléments impactés par les TAGs recherchés (3). Le résultat s’affiche dans l’IHM en spécifiant le type de relations entre les TAGs ainsi que toutes les informations disponibles sur les éléments impactés.

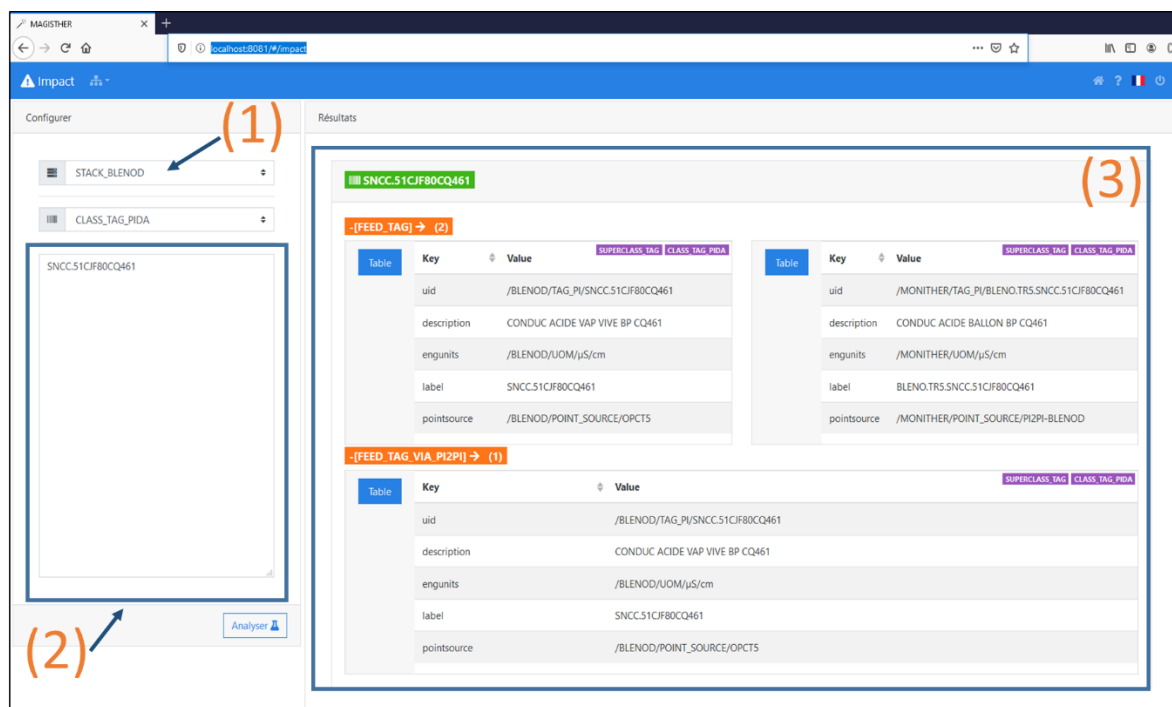


Figure 63 Maquette du module d’analyse d’impact

Ce module est encore à l’état de test et ne permet que de requêter sur l’environnement de Blenod. Dans les prochains mois, nous allons devoir adapter la maquette afin de pouvoir requêter sur l’ensemble de la cartographie. De plus, nous allons organiser des ateliers afin que les utilisateurs finaux puissent utiliser l’outil et nous faire un retour sur son ergonomie et les fonctionnalités qui pourraient être utiles à l’analyse d’impact.

## 7.2.4 Plan du SI

Dans chaque entreprise, on dessine des diagrammes de réseaux complexes afin de représenter schématiquement l'architecture du SI.

Par exemple, ces entreprises créent ces diagrammes via des applications comme Microsoft Visio.

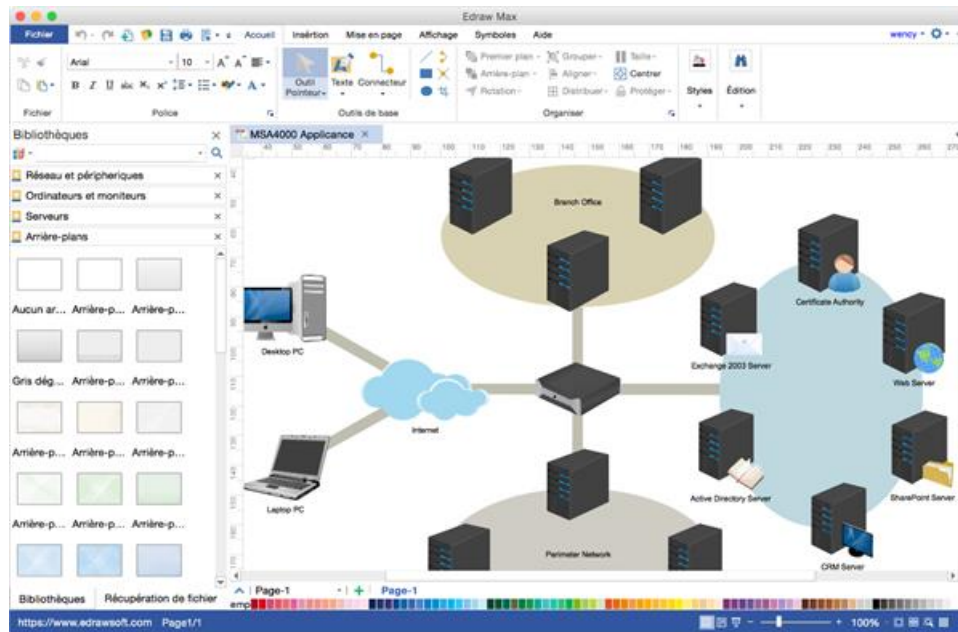


Figure 64 Exemple de représentation via Visio<sup>18</sup>

Cette représentation nécessite un réel effort de design afin de représenter, machine par machine l'infrastructure de l'entreprise. En plus de la complexité à mettre en place, cette représentation n'est représentative de la réalité qu'à l'instant où elle est dessinée. Dès lors qu'un système est ajouté, supprimé ou modifié, la carte devient obsolète et doit être reprise.

Nous avons découvert une technologie open source permettant de générer des diagrammes en fonction d'un fichier texte en entrée. Cette particularité fait de cette solution, une méthode beaucoup moins complexe afin de générer des diagrammes (pas d'apprentissage de logiciel ou de langage particulier). De plus, le diagramme généré en sortie est en format SVG afin d'être implanté facilement dans un navigateur web et donc dans un module de notre portail.

Aujourd'hui, notre base de connaissances contient les informations de l'infrastructure réseaux du SI. Le fichier texte permettant de générer ce diagramme pourrait donc être généré à l'aide d'une requête Cypher dans notre base de graphes. Ainsi, nous aurions une représentation du SI

<sup>18</sup> Source : <https://www.microsoft.com/fr-fr/microsoft-365/visio/flowchart-software>

simple à mettre place, toujours représentative de la réalité (car produit par le graphe) et accessible via le portail Web de MAGISTERHER.

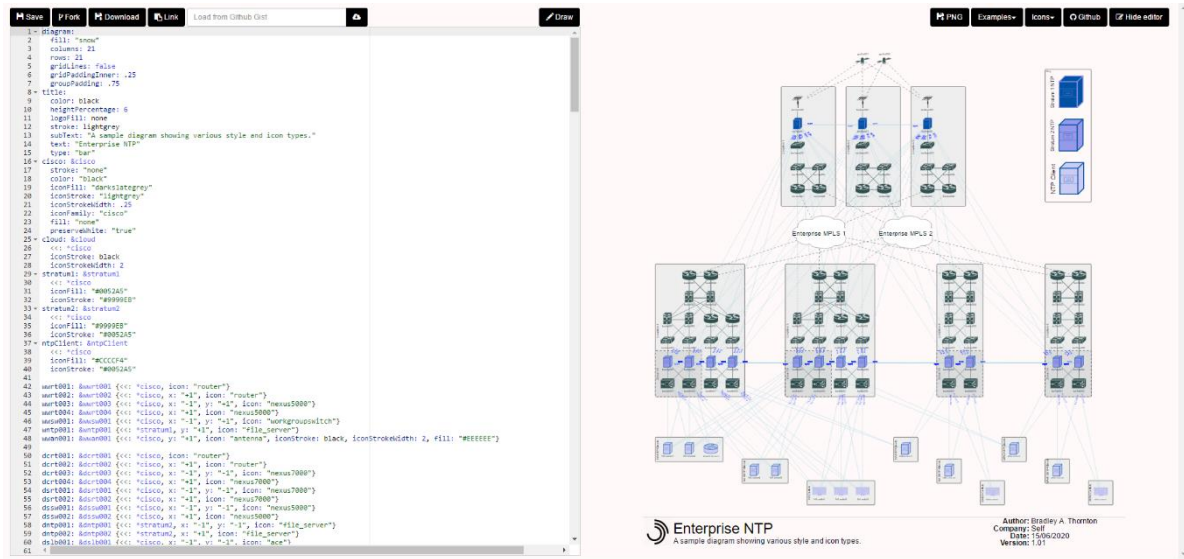


Figure 65 Exemple de représentation d'un SI via un fichier texte

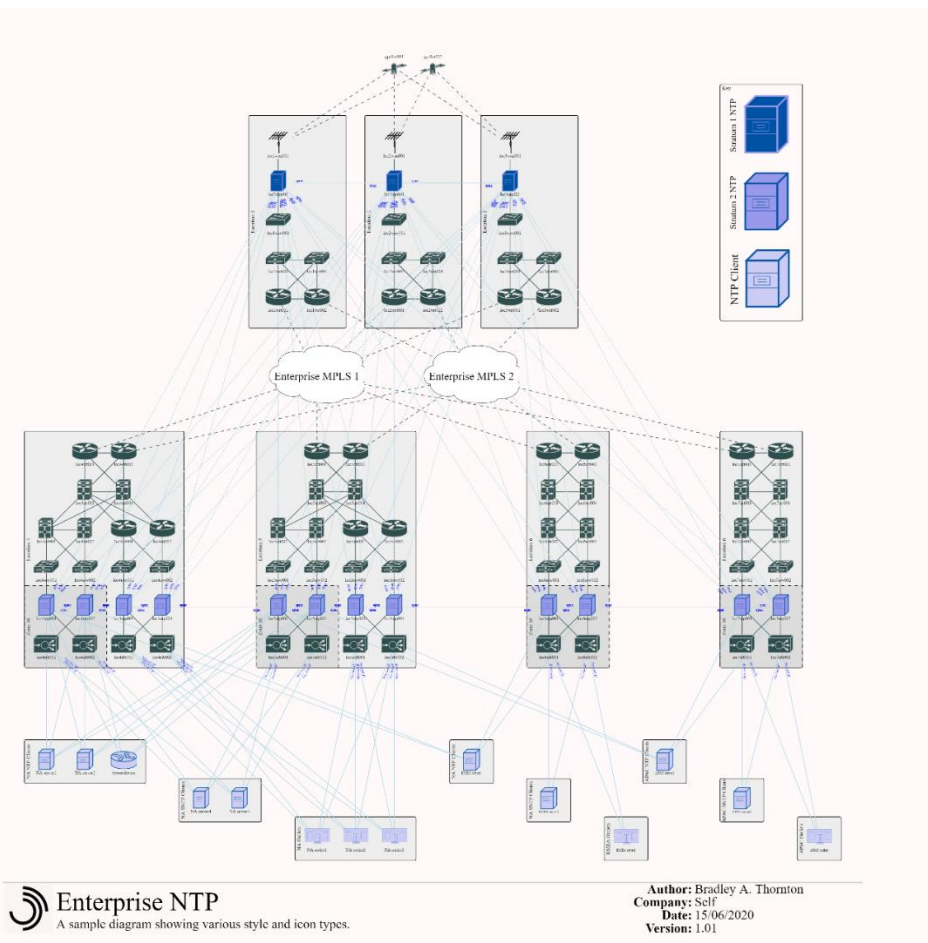


Figure 56 Exemple de représentation d'un SI via l'outil drawthe.net<sup>19</sup>

<sup>19</sup> Source : <https://github.com/cidrblock/drawthe.net>

## 8 CONCLUSION

---

MAGISTHER est une démarche pérenne en amélioration continue.

Des ateliers vont être organisés dans les prochains mois afin de recueillir l'avis des utilisateurs finaux sur les différents modules et de faire les adaptations qui en découleront. Comme évoqué dans le paragraphe précédent, beaucoup de nouvelles fonctionnalités et modules vont voir le jour très prochainement amenant de nouvelles phases de développement, de tests et d'ateliers.

En guise de conclusion, je vais faire un point sur le projet MAGISTHER et mon implication dans son développement mais également sur mes trois années d'alternance au sein d'EDF Ingeum ainsi que mon futur dans l'ingénierie.

### 8.1 BILAN PROFESSIONNEL

Au cours de ces trois dernières années, j'ai pu étoffer mon éventail de compétences en travaillant sur le projet MAGISTHER.

J'ai appris à développer des applications web en développant le module de supervision dans le portail web du projet. J'ai appris à me servir des technologies d'ETL afin de traiter et restructurer les données pour les adapter pour la base de graphes. J'ai également pu découvrir les bases de données orientées graphe afin de créer la cartographie des données du SI.

Au cours de ce projet, j'ai dû piloter à plusieurs reprises des prestataires, ce qui m'a permis de travailler mes qualités relationnelles et mes facultés d'encadrement. L'écriture de documentation ainsi que les présentations du projet m'ont appris à être plus à l'aise à l'oral et à rédiger des documentations techniques.

### 8.2 BILAN PERSONNEL

Etant d'un naturel timide, j'avais beaucoup de mal à parler aux autres et j'essayais souvent de résoudre mes problèmes seul plutôt que de « déranger » quelqu'un. Mais en travaillant en équipe, j'ai compris que l'entraide était l'un des points-clés du bon fonctionnement de cette dernière. J'ai appris à aller plus vers les autres, à m'intéresser aux autres projets sur lesquels ils travaillaient.

J'ai pu collaborer avec des personnes en dehors de notre unité pour des projets annexes comme le montage de la vidéo de départ en retraite de notre directeur.

Au cours de cette alternance, j'ai appris à me dépasser sur le plan relationnel et technique et cela m'a permis de devenir un meilleur ingénieur.

### 8.3 PERSPECTIVE D'AVENIR

En entrant à EDF Ingeum (anciennement Centre d'Ingénierie Thermique d'EDF) j'étais tiraillé entre plusieurs domaines de l'ingénierie qui me plaisaient beaucoup. Je sortais d'un IUT Génie Electrique et Informatique Industrielle qui m'avais permis de découvrir de nombreux domaines dont : L'automatisme Industrielle et l'Informatique Industrielle.

J'ai ensuite réalisé un stage dans l'automatisme car j'étais vraiment passionné par les automates et la programmation de ces derniers.

Lorsque j'ai postulé pour l'alternance chez EDF, je savais très bien que je ne travaillerais pas sur ce genre de technologie. Néanmoins, j'ai accepté le poste car j'aimais l'informatique et je voulais avoir une expérience dans mes deux domaines de prédilection afin de pouvoir, au terme de mon alternance choisir la branche qui me plairait le plus.

Lorsque j'ai commencé à travailler sur le projet MAGISTHER, j'ai compris que l'Automatisme et l'Informatique Industrielle pouvaient être liés. En effet, si on assimile le SI à une usine, la somme des systèmes qui le compose est un automate qu'il faut programmer et orchestrer. La conception, la récupération de données ... tout est automatisé afin de générer la carte et le portail de MAGISTHER. Je me suis retrouvé, dans ce projet, à la croisée des domaines que j'appréciais le plus. Par ailleurs, l'analyse et la restructuration des données ainsi que les bases de graphes m'ont particulièrement intéressé, que ce soit la conception de scripts afin de valoriser les données mais aussi la réalisation de la structure du graphe afin de construire la cartographie des données.

Je souhaite donc continuer à travailler dans le domaine de la gouvernance de données, afin d'en apprendre plus sur les technologies d'ETL et de bases de graphes.

J'ai eu la chance de me voir proposé une offre d'emploi par EDF Ingeum que j'ai accepté, afin de continuer l'extension du périmètre de MAGISTHER, l'étude de la supervision et d'autres projets touchants à la gouvernance du SI.

## 9 TABLE DES FIGURES

---

|  |    |
|--|----|
| Figure 1 Les deux grands métiers d'EDF.....  | 6  |
| Figure 2 Chaîne de valeur du secteur de l'énergie .....  | 7  |
| Figure 3 Courbe représentative de la consommation d'électricité moyenne en France pour une journée.....                    | 8  |
| Figure 4 Production Française d'électricité en 2017.....   | 8  |
| Figure 5 Organigramme de la DTEAM .....  | 10 |
| Figure 6 Organigramme EDF Ingeum 2018.....   | 10 |
| Figure 7 Stratégie EDF Ingeum Ligne Internationale.....  | 11 |
| Figure 8 Industrie 4.0 : la 4 <sup>ème</sup> révolution Industrielle .....   | 13 |
| Figure 9 L'ingénierie au cœur de la gestion de l'information sur l'ensemble du cycle de vie d'un actif de production ..... | 14 |
| Figure 10 La chaîne d'approvisionnement des données .....  | 15 |
| Figure 11 Les problèmes liés à la chaîne d'approvisionnement des données .....   | 16 |
| Figure 12 Notion de TAG .....  | 17 |
| Figure 13 Mapping de tag .....   | 17 |
| Figure 14 Exemple de reconstitution de la chaîne d'acheminement des données .....  | 17 |
| Figure 15 Magisther et la chaîne d'approvisionnement .....   | 18 |
| Figure 16 : Image provenant du document officiel de la présentation du cap 2030 .....                                      | 19 |
| Figure 17 Le logiciel Waze .....   | 20 |
| Figure 18 Urbanisation du SI .....   | 21 |
| Figure 19 Le rocher de Sisyphe : de la mythologie à une vision à l'ère moderne .....                                       | 22 |
| Figure 20 Gestion des connaissances du SI .....  | 23 |
| Figure 21 Chaîne de valeur inhérente à MAGISTHER .....   | 24 |
| Figure 22 Magisther et la chaîne d'approvisionnement .....   | 25 |
| Figure 23 Exemple de lien entre deux identifiants indirect .....   | 25 |
| Figure 24 Les composantes du système d'information.....  | 26 |
| Figure 25 Analogie entre les données et les étoiles .....  | 26 |
| Figure 26 Le SI sous forme de couche caractérise en son sein des objets eux même appartenant à des référentiels.....       | 27 |
| Figure 27 Vision macroscopique du fonctionnement de la plateforme MAGISTHER .....  | 28 |
| Figure 28 Module d'accès aux catalogues de données des sites .....   | 29 |
| Figure 29 Module d'accès aux catalogues ciblant un site : CETAC.....   | 29 |
| Figure 30 Module de cartographie illustrant les données liées .....  | 30 |
| Figure 31 Module de cartographie présentant les données d'un nœud .....  | 30 |
| Figure 32 Vision macroscopique du fonctionnement de la plateforme MAGISTHER .....  | 31 |
| Figure 33 Illustration du fonctionnement de GITLAB .....   | 32 |
| Figure 34 Fonctionnement interne de MAGISTHER .....  | 33 |
| Figure 35 Vision macroscopique du fonctionnement de la plateforme MAGISTHER (détaillée) .....                              | 34 |
| Figure 36 Exemple d'utilisation des requêtes Cypher .....  | 35 |
| Figure 37 Intégration de la cartographie des données dans notre portail.....   | 35 |
| Figure 38 Exemple de représentation des données dans le module de la cartographie .....                                    | 36 |
| Figure 39 Exemple de lien entre différentes instances de TAG.....  | 36 |

|  |    |
|--|----|
| Figure 40 Exemple de lien avec la source du TAG.....   | 37 |
| Figure 41 Magisther et la chaîne d’approvisionnement .....   | 37 |
| Figure 42 Vision macroscopique du fonctionnement de la plateforme MAGISTHER .....                    | 38 |
| Figure 43 Les couches du SI et l’extension de MAGISTHER .....  | 39 |
| Figure 44 Use case de SFR sur la gouvernance de l’information .....                                  | 40 |
| Figure 45 Les concepts ajoutés à la cartographie.....  | 40 |
| Figure 46 Catalogue des données des serveurs .....   | 41 |
| Figure 47 Implantation des données des serveurs dans le module cartographie existant .....           | 42 |
| Figure 48 Ecran de supervision EDF Ingeum .....  | 43 |
| Figure 49 Schématisation de la génération des écrans de supervision.....                             | 44 |
| Figure 50 Représentation graphique des serveurs.....   | 45 |
| Figure 51 Panel d’information sur un serveur.....  | 45 |
| Figure 52 Représentation synthétique des serveurs du SI.....   | 46 |
| Figure 53 Représentation synthétique des serveurs du SI (Filtrage).....                              | 46 |
| Figure 54 Représentation des serveurs sous forme de tableau téléchargeable dans au format Excel..... | 47 |
| Figure 55 Module de supervision version mobile.....  | 47 |
| Figure 56 Système de génération composite de document par Samuel PARFOURU et Pascal NEGROS .....     | 49 |
| Figure 57 Cartographie du SI dans MAGISTHER .....  | 49 |
| Figure 58IHM de visualisation de document .....  | 50 |
| Figure 59 IHM de visualisation de document centré sur la documentation.....                          | 51 |
| Figure 60 Principe de Génération automatique de document .....                                       | 52 |
| Figure 61 (Génération d’un document basé sur les éléments du graphe) .....                           | 53 |
| Figure 62 IHM exploitant DagreJS répondant au besoin de supervision d’un script .....                | 56 |
| Figure 63 Maquette du module d’analyse d’impact.....   | 57 |
| Figure 64 Exemple de représentation via Visio .....  | 58 |
| Figure 65 Exemple de représentation d’un SI via un fichier texte.....                                | 59 |