

ENQUÊTE MÉTIER

Auteur : Mathéo Carval

FISA A3 INFORMATIQUE

CESI Montpellier

Table des matières

1. INTRODUCTION GÉNÉRALE	4
2. ENQUÊTE MÉTIER : LE DATA ENGINEER EN 2030	6
2.1. Analyse prospective et positionnement stratégique	6
2.2. Le Data Engineer au cœur des transformations	6
2.3. Analyse du Marché de l'Emploi et des Compétences	7
2.4. Synthèse des Entretiens Professionnels.....	9
Question 1 : Comment définissez-vous votre rôle aujourd'hui et comment s'articule-t-il avec les autres services ?	10
Question 2 : Quelles sont les compétences techniques et les soft skills les plus critiques pour réussir dans ce métier ?	11
Question 3 : Comment percevez-vous l'évolution du métier d'ici 5 à 10 ans avec la montée en puissance de l'IA et de l'automatisation ?	12
Question 4 : Comment les enjeux de développement durable impactent-ils votre manière de travailler et vos choix techniques ?	13
Question 5 : Quels sont les principaux défis que vous rencontrez au quotidien et comment les surmontez-vous ?	14
Question 6 : Quelles sont les évolutions technologiques majeures que vous avez observées ces dernières années et comment vous êtes-vous adapté ?	15
Question 7 : Quels conseils donneriez-vous à quelqu'un qui débute dans le métier de Data Engineer ?.....	16
Question 8 : Comment gérez-vous l'équilibre entre innovation technologique et maintien de la stabilité des systèmes en production ?	16
Question 9 : Quel est votre rapport au travail en équipe et comment collaborez-vous avec les autres profils data (Data Analysts, Data Scientists, etc.) ?	17
Question 10 : Quelles sont les erreurs les plus fréquentes que vous voyez chez les Data Engineers juniors et comment les éviter ?	18
Conclusion générale des entretiens	19
3. ENQUÊTE SUR SOI - ANALYSE INTROSPECTIVE	20
3.1. Feedback 360° : Le regard de l'entourage	20
3.2. Profil de personnalité et soft skills	21
3.3. Système de valeurs et motivations profondes.....	23

4. LIEN ENTRE ENQUÊTE METIER ET ENQUÊTE SUR SOI.....	25
4. CIBLE MÉTIER	26
Définition du poste et missions principales	26
Lien avec les Objectifs de Développement Durable.....	26
Vision à 5-10 ans	27
5. BIBLIOGRAPHIE ET SOURCES	28
Glossaire.....	30

1. INTRODUCTION GÉNÉRALE

Le métier d'ingénieur connaît une transformation profonde à l'aube de 2030. Qu'il s'agisse d'ingénieurs en génie civil, en mécanique, en électronique ou en informatique, tous font face à des défis communs : la digitalisation des processus, l'émergence de l'intelligence artificielle, et la nécessité d'intégrer la dimension environnementale dans leurs pratiques. L'ingénieur de demain ne sera plus seulement un expert technique résolvant des problèmes isolés ; il devra devenir un acteur transverse capable de naviguer entre innovation technologique, contraintes éthiques et responsabilité sociétale.

Dans le domaine spécifique de l'informatique, et plus particulièrement dans celui de la donnée, cette mutation est particulièrement visible. Le métier de Data Engineer incarne parfaitement cette évolution : portée par l'explosion des volumes de données et l'émergence de l'intelligence artificielle générative, cette profession se transforme radicalement. À l'horizon 2030, le Data Engineer ne sera plus seulement un architecte de pipelines de données ; il deviendra un garant de la qualité, de la gouvernance et de l'éthique de l'information au sein des organisations.

Ce dossier d'enquête métier s'inscrit dans le cadre de ma formation d'ingénieur en informatique à CESI Montpellier, en alternance chez Hérault Logement. Cette démarche prospective vise à confronter ma vision actuelle du métier de Data Engineer avec les réalités d'un secteur en pleine mutation, où la donnée est devenue l'actif stratégique majeur des entreprises, au même titre que les ressources financières ou humaines.

Mon parcours m'a naturellement orienté vers ce domaine. Dès le lycée, j'ai développé un intérêt marqué pour l'informatique et la résolution de problèmes logiques. Cet intérêt s'est confirmé durant mes années de BUT Informatique à l'IUT de Valence. Cette formation pluridisciplinaire m'a permis de découvrir les différentes facettes du métier d'informaticien : développement logiciel, administration système, réseaux, bases de données et gestion de projet. C'est particulièrement le domaine de la donnée qui a retenu mon attention lors des cours et projets de bases de données. L'aspect technique de la gestion des données, combiné à leur dimension stratégique pour les organisations, m'a convaincu d'orienter mon projet professionnel vers le métier de Data Engineer.

Aujourd'hui, en alternance chez Hérault Logement en tant qu'administrateur système et data analyst, je suis au cœur des problématiques de gestion de l'information. Mon rôle m'amène à travailler sur plusieurs dimensions complémentaires : la gestion de

l'infrastructure Active Directory qui structure l'identité numérique de l'entreprise, le développement de dashboards Power BI pour l'analyse des données GLPI qui permettent d'optimiser la gestion du parc informatique, et plus globalement la maintenance de l'ensemble de l'écosystème IT de l'entreprise. Cette expérience concrète me permet d'appréhender les enjeux réels du métier et de comprendre comment la donnée circule, se transforme et génère de la valeur au sein d'une organisation de taille moyenne comme Hérault Logement.

L'objectif de ce travail est d'abord de décrypter en profondeur les compétences techniques, les responsabilités et le positionnement stratégique qui définiront les ingénieurs et Data Engineer de demain, dans un contexte où l'IA et l'automatisation redéfinissent les contours du métier. De plus, cette enquête constitue une démarche introspective rigoureuse visant à évaluer l'adéquation entre mon profil actuel, mes valeurs professionnelles et les exigences du métier ciblé.

2. ENQUÊTE MÉTIER : LE DATA ENGINEER EN 2030

2.1. Analyse prospective et positionnement stratégique

Le Data Engineer de 2030 se définit par son positionnement à l'intersection de plusieurs domaines d'expertise : l'architecture de données, le développement logiciel, l'administration système et la compréhension des enjeux métier. Contrairement aux années 2010-2020 où le métier consistait principalement à construire des pipelines ETL* pour alimenter des entrepôts de données, le Data Engineer moderne doit désormais concevoir des architectures data complexes capables de gérer des flux en temps réel, d'intégrer des modèles d'intelligence artificielle et de garantir la conformité réglementaire.

Cette évolution transforme radicalement la définition de fonction. Le Data Engineer devient un architecte de l'information qui doit anticiper les besoins futurs de l'organisation, concevoir des systèmes évolutifs et garantir la qualité des données de bout en bout. Il n'est plus un simple exécutant technique mais un conseiller stratégique capable d'évaluer l'impact à long terme des choix technologiques sur la performance de l'entreprise.

Le positionnement du Data Engineer dans l'organigramme évolue également. Traditionnellement rattaché aux équipes IT, il se rapproche aujourd'hui des directions métier et des équipes data science pour former des cellules transverses dédiées à la valorisation de la donnée. Cette proximité avec le business permet une meilleure compréhension des cas d'usage et une adaptation plus rapide aux évolutions des besoins.

L'analyse des offres d'emploi sur les cinq dernières années montre une nette progression de la demande pour ce profil. Les entreprises recherchent des Data Engineers capables de maîtriser à la fois les technologies cloud* (AWS, Azure, GCP), les outils de big data* (Spark, Kafka, Airflow) et les problématiques de gouvernance des données. La pénurie de compétences dans ce domaine crée une situation favorable pour les jeunes ingénieurs, avec des perspectives salariales attractives et des possibilités d'évolution rapide.

2.2. Le Data Engineer au cœur des transformations

Au sein des organisations modernes, le Data Engineer joue un rôle central dans trois transformations majeures : la transformation numérique, la transformation data-driven et la transformation écologique.

La transformation numérique repose en grande partie sur la capacité des entreprises à exploiter leurs données. Le Data Engineer est le garant de cette exploitation en concevant les infrastructures qui permettent de collecter, stocker et traiter l'information. Dans mon expérience actuelle chez Hérault Logement, cette dimension est particulièrement visible : la mise en place de dashboards Power BI pour analyser les données GLPI permet aux équipes de prendre des décisions éclairées sur la gestion du parc informatique et l'optimisation des ressources.

La transformation data-driven consiste à faire de la donnée le socle des décisions stratégiques. Le Data Engineer facilite cette transformation en garantissant que les données sont fiables, accessibles et compréhensibles pour les équipes métier. Il doit construire des pipelines robustes qui alimentent les outils d'analyse et de visualisation, tout en assurant la traçabilité et la documentation des transformations appliquées. Cette responsabilité implique une rigueur technique importante mais également une capacité à dialoguer avec les utilisateurs finaux pour comprendre leurs besoins réels.

La transformation écologique est désormais indissociable du métier de Data Engineer. À l'horizon 2030, la sobriété numérique* devient un impératif. Le Data Engineer doit concevoir des architectures qui optimisent la consommation énergétique, limitent le stockage inutile et privilégient le traitement efficace des données. Cette dimension écologique se traduit par des choix techniques concrets : privilégier le streaming* plutôt que le batch* quand c'est pertinent, mettre en place des politiques de rétention des données, optimiser les requêtes pour réduire la sollicitation des serveurs.

Dans le domaine spécifique du cloud computing*, le Data Engineer doit également être vigilant sur les coûts d'infrastructure. Une architecture mal conçue peut générer des dépenses importantes tout en ayant un impact environnemental significatif. La capacité à dimensionner correctement les ressources, à mettre en place des mécanismes d'auto-scaling* et à monitorer les performances devient donc une compétence clé.

2.3. Analyse du Marché de l'Emploi et des Compétences

Le marché de l'emploi pour les Data Engineers connaît une dynamique exceptionnelle depuis plusieurs années. Les études récentes mettent en évidence plusieurs tendances structurantes pour la décennie 2020-2030.

Une demande en forte croissance : Selon les données de l'APEC et de France Travail, les offres d'emploi pour les profils Data Engineer ont progressé de plus de 300% entre 2018

et 2024. Cette croissance s'explique par la généralisation des stratégies data au sein des entreprises, qui considèrent désormais la donnée comme un actif stratégique majeur. Les secteurs les plus demandeurs sont traditionnellement la finance, les assurances, le e-commerce et les télécommunications, mais on observe une diversification vers tous les secteurs d'activité, y compris l'industrie, la santé et le secteur public.

Une pénurie de compétences qualifiées : Malgré l'augmentation des formations en data engineering, l'offre de profils qualifiés peine à suivre la demande. Cette situation crée un marché favorable aux jeunes diplômés, avec des salaires d'embauche attractifs. Selon l'enquête de l'IESF (Ingénieurs et Scientifiques de France), un Data Engineer junior peut prétendre à un salaire annuel brut compris entre 38K€ et 45K€, tandis qu'un profil confirmé (5-10 ans d'expérience) peut atteindre 60K€ à 80K€, voire davantage dans les grandes métropoles ou les entreprises du secteur tech.

L'évolution des compétences attendues : L'analyse des offres d'emploi révèle une évolution significative des compétences recherchées. Les cinq dernières années ont vu émerger de nouvelles exigences :

- Maîtrise du cloud : AWS, Azure et GCP sont devenus incontournables. Les entreprises migrent massivement leurs infrastructures vers le cloud, et le Data Engineer doit être capable de concevoir et opérer des architectures cloud-native*.
- Compétences en temps réel : Les technologies de streaming (Kafka, Flink, Spark Streaming) gagnent en importance face aux traitements batch traditionnels. La capacité à traiter des flux de données en temps réel devient discriminante.
- Gouvernance et qualité des données : Au-delà de la construction de pipelines, le Data Engineer doit maîtriser les outils et les pratiques de data governance*, de data quality* et de data lineage*. La conformité RGPD* et la traçabilité des données sont devenues des enjeux majeurs.
- Infrastructure as Code (IaC) : La capacité à automatiser le déploiement et la gestion des infrastructures via du code (Terraform*, Ansible*) est désormais attendue. Cette compétence permet d'améliorer la reproductibilité, la fiabilité et la scalabilité* des systèmes.

- DataOps et MLOps : L'intégration des pratiques DevOps* au monde de la donnée et du machine learning devient standard. Le Data Engineer doit savoir mettre en place des pipelines CI/CD*, des tests automatisés et des mécanismes de monitoring*.

L'impact de l'intelligence artificielle : L'émergence de l'IA générative et des modèles de machine learning complexes redéfinit partiellement le métier. Le Data Engineer doit désormais collaborer étroitement avec les Data Scientists et les ML Engineers pour construire des pipelines qui alimentent et servent les modèles d'IA. Cette évolution nécessite une compréhension minimale du machine learning et des problématiques spécifiques à l'entraînement et au déploiement de modèles.

Les soft skills de plus en plus valorisées : Au-delà des compétences techniques, les recruteurs recherchent des profils capables de communiquer efficacement avec les équipes métier, de travailler en mode agile* et de faire preuve d'autonomie dans la résolution de problèmes complexes. La capacité à documenter son travail, à partager ses connaissances et à collaborer dans des environnements multiculturels devient déterminante, notamment dans les grandes entreprises internationales.

Perspectives à 10 ans : Les projections indiquent que la demande continuera de croître jusqu'en 2030, avec une légère inflexion possible liée à l'automatisation de certaines tâches par l'IA. Toutefois, l'augmentation exponentielle des volumes de données et la complexification des architectures devraient maintenir une forte tension sur le marché. Les Data Engineers capables de démontrer leur valeur ajoutée sur les enjeux de gouvernance, d'optimisation et de durabilité auront les meilleures perspectives d'évolution, tant en termes de responsabilités que de rémunération.

2.4. Synthèse des Entretiens Professionnels

Pour approfondir ma compréhension du métier, j'ai mené une série d'entretiens auprès de quatre professionnels occupant des fonctions liées au domaine de la data. Ces échanges m'ont permis de confronter ma vision théorique du métier avec la réalité du terrain, d'identifier les compétences réellement mobilisées au quotidien et de mieux comprendre les défis auxquels sont confrontés les Data Engineers en 2025

Profil des personnes interrogées

Entretien n°1 : Un collègue senior de mon entreprise d'alternance (8 ans d'expérience en infrastructure IT et gestion de données), responsable de l'architecture des systèmes d'information et référent technique sur les projets de migration cloud

Entretien n°2 : Un membre de ma famille travaillant comme Architecte Data (12 ans d'expérience) dans le secteur bancaire, expert en gouvernance des données et conformité RGPD

Entretien n°3 : Une ancienne camarade de BUT Informatique (promotion 2022), actuellement DataOps Engineer (3 ans d'expérience) dans une startup tech à Lyon, spécialisée dans l'automatisation des pipelines de données

Entretien n°4 : Une connaissance du secteur travaillant comme Manager d'équipe Data (10 ans d'expérience) dans une entreprise du secteur de l'assurance, responsable d'une équipe de 6 Data Engineers

Question 1 : Comment définissez-vous votre rôle aujourd'hui et comment s'articule-t-il avec les autres services ?

L'ensemble des personnes interrogées s'accordent sur le caractère transverse du métier de Data Engineer. Contrairement à d'autres fonctions IT plus cloisonnées, le Data Engineer se positionne à l'interface entre plusieurs univers : les équipes infrastructure, les développeurs applicatifs, les data scientists et les équipes métier. Cette position centrale implique une capacité à parler plusieurs "langages" et à traduire les besoins métier en solutions techniques.

Mon collègue senior chez Hérault Logement décrit son rôle comme celui d'un "facilitateur de la valorisation de la donnée". Il insiste sur l'importance de construire des fondations solides qui permettent ensuite aux équipes métier de travailler efficacement. Cette dimension de "service aux autres équipes" est récurrente dans les témoignages.

Le membre de ma famille travaillant comme architecte Data dans le secteur bancaire, quant à lui, met l'accent sur la dimension stratégique de la fonction. Au-delà de l'exécution technique, il participe activement aux décisions d'architecture globale du système d'information et conseille la direction sur les investissements technologiques à réaliser. Cette évolution vers un rôle de conseil confirme la tendance observée dans l'analyse prospective.

Ces retours confirment que le Data Engineer ne peut pas se contenter d'être un expert technique isolé. La dimension relationnelle et la capacité à comprendre les enjeux business sont tout aussi importantes que la maîtrise des technologies. Dans mon expérience actuelle chez Hérault Logement, je constate déjà cette nécessité de dialoguer avec différents services pour comprendre leurs besoins en matière de données et de reporting. Développer cette compétence de "traducteur" entre le technique et le métier sera un enjeu majeur pour moi.

Question 2 : Quelles sont les compétences techniques et les soft skills les plus critiques pour réussir dans ce métier ?

Sur le plan technique, trois grandes catégories de compétences ressortent systématiquement :

- Maîtrise des technologies de traitement de données : SQL avancé, Python (parfois Scala), connaissance approfondie d'au moins un moteur de traitement distribué (Spark en tête), maîtrise des outils d'orchestration (Airflow, Dagster).
- Compétences cloud et infrastructure : Connaissance d'au moins un provider cloud majeur (AWS domine, suivi d'Azure), maîtrise de l'Infrastructure as Code, compréhension des principes de réseau et de sécurité.
- Pratiques DevOps et automatisation : Git, CI/CD, conteneurisation (Docker, Kubernetes), monitoring et observabilité.

Concernant les soft skills, la rigueur est citée par tous les interviewés comme la compétence humaine la plus critique. Mon contact architecte Data insiste particulièrement sur ce point : "Dans notre métier, une petite erreur peut avoir des conséquences graves sur les données de production. Il faut toujours vérifier deux fois, documenter ce qu'on fait et ne jamais prendre de raccourcis." Cette exigence de rigueur rejoint directement ma propre approche du travail.

La communication arrive en seconde position. La connaissance du secteur travaillant en management souligne que les meilleurs profils techniques ne sont pas toujours les plus efficaces s'ils ne savent pas expliquer ce qu'ils font et pourquoi ils le font. Elle insiste sur l'importance de la documentation, des compte-rendus clairs et de la capacité à vulgariser les concepts techniques.

L'autonomie et la capacité de résolution de problèmes sont également largement mentionnées. Le Data Engineer est souvent confronté à des situations inédites où les solutions ne sont pas documentées. Savoir défricher, tester des hypothèses et persévérer face à des problèmes complexes est essentiel.

Ma rigueur naturelle et mon autonomie dans le travail correspondent aux attentes du métier. En revanche, je dois progresser sur la dimension communication, notamment sur ma capacité à partager proactivement l'avancement de mes travaux et à solliciter de l'aide quand c'est nécessaire. Mon côté réservé peut parfois être perçu comme un manque de transparence, ce qui est un point d'attention important pour un métier aussi collaboratif.

Question 3 : Comment percevez-vous l'évolution du métier d'ici 5 à 10 ans avec la montée en puissance de l'IA et de l'automatisation ?

Cette question a suscité des réponses nuancées. Aucun des interviewés ne pense que l'IA va remplacer le Data Engineer à court ou moyen terme, mais tous reconnaissent que le métier va significativement évoluer.

Mon ancienne camarade de BUT travaillant comme DataOps Engineer est la plus optimiste : "L'IA va nous débarrasser des tâches répétitives et nous permettre de nous concentrer sur ce qui a vraiment de la valeur : l'architecture, la qualité des données, la compréhension des besoins métier." Elle voit l'IA comme un assistant qui accélère le développement, aide au debugging et propose des optimisations, mais qui ne remplace pas le jugement et l'expertise humaine.

L'architecte Data de ma famille adopte une position plus prudente : "Certaines tâches standardisées vont effectivement être automatisées. Les pipelines ETL simples, la génération de code boilerplate, les premiers niveaux de diagnostic d'erreurs... Tout ça peut être fait par l'IA. Mais la vraie difficulté dans notre métier, c'est de comprendre les données, de garantir leur qualité et de concevoir des architectures robustes. Ça, l'IA ne sait pas le faire toute seule."

Mon collègue senior partage cette analyse : "L'IA va changer nos outils et nos méthodes, mais pas fondamentalement notre rôle. Je pense que dans 10 ans, on passera beaucoup moins de temps à écrire du code et beaucoup plus de temps à le reviewer, à valider la qualité et à s'assurer que tout fonctionne comme prévu. Le métier va devenir plus intellectuel et moins manuel."

Ma connaissance manager ajoute une dimension organisationnelle : "L'IA va aussi créer de nouveaux besoins. Il faudra construire des pipelines pour alimenter les modèles, gérer les versions des données d'entraînement, monitorer les performances... Le métier va se complexifier sur ces aspects."

Ces perspectives sur l'évolution du métier correspondent à ma propre vision. L'IA ne va pas supprimer le métier mais le transformer. Cette transformation sera plutôt positive si on sait s'adapter et monter en compétences sur les aspects les plus stratégiques. Mon objectif est de ne pas me cantonner aux tâches d'exécution qui risquent d'être automatisées, mais de développer dès maintenant une expertise sur les sujets d'architecture, de qualité et de gouvernance qui resteront du domaine de l'humain.

Question 4 : Comment les enjeux de développement durable impactent-ils votre manière de travailler et vos choix techniques ?

Mon collègue senior explique avoir mis en place des mécanismes de monitoring de la consommation énergétique : "On a réalisé qu'on faisait tourner des traitements surdimensionnés qui consommaient inutilement des ressources. En optimisant nos requêtes et en dimensionnant mieux nos infrastructures, on a divisé par 3 notre facture cloud, ce qui a aussi un impact écologique direct."

L'architecte Data a une approche plus stratégique : "On a mis en place une politique de rétention des données beaucoup plus stricte. Avant, on gardait tout 'au cas où'. Maintenant, on se pose systématiquement la question : est-ce que ces données ont vraiment de la valeur ? Si ce n'est pas le cas, on les supprime. Ça réduit les coûts de stockage et l'empreinte carbone."

Mon ancienne camarade DataOps Engineer insiste sur l'importance du monitoring : "On a implémenté des alertes sur la consommation des ressources. Si un job consomme plus que prévu, on est notifié immédiatement et on peut investiguer. Souvent, c'est signe d'un problème de performance qu'on peut corriger."

La connaissance manager évoque également la dimension organisationnelle : "On sensibilise toute l'équipe à ces enjeux. Chaque revue de code inclut maintenant une vérification de l'optimisation. C'est devenu une partie intégrante de notre culture."

La sobriété numérique est un sujet qui me tient à cœur et qui rejoint mes valeurs personnelles. Dans mes futures fonctions, je souhaite intégrer systématiquement

cette dimension dès la phase de conception des architectures. L'optimisation des ressources n'est pas seulement une question de coût, c'est aussi une responsabilité écologique. Mon expérience actuelle sur l'optimisation de l'infrastructure chez Hérault Logement me prépare déjà à ces enjeux.

Question 5 : Quels sont les principaux défis que vous rencontrez au quotidien et comment les surmontez-vous ?

La qualité des données est le défi le plus souvent cité. L'architecte Data explique : "On passe une grande partie de notre temps à comprendre pourquoi les données ne sont pas conformes aux attentes, à investiguer les sources de problèmes et à mettre en place des contrôles. C'est un travail de fond qui n'est jamais terminé." Ce retour confirme l'importance de la rigueur et de la persévérance dans le métier.

La gestion de la complexité technique est également un défi majeur. Mon collègue senior décrit : "Les écosystèmes data sont devenus incroyablement complexes. On jongle entre 15 technologies différentes, chacune avec ses spécificités. Il faut constamment se former, tester, documenter. La charge mentale est importante." Cette complexité nécessite une organisation rigoureuse et une bonne capacité d'apprentissage continu.

La communication avec les équipes métier représente un défi relationnel important. La connaissance manager souligne : "Il y a souvent un décalage entre ce que les métiers pensent être facile à faire et la réalité technique. Notre rôle est de gérer ces attentes, d'expliquer les contraintes et de trouver des compromis." Cette dimension confirme l'importance des soft skills.

L'équilibre entre vitesse et qualité est un dilemme récurrent. Mon ancienne camarade DataOps Engineer explique : "On nous demande de livrer vite, mais on ne peut pas sacrifier la qualité. Il faut trouver le bon équilibre, savoir dire non parfois, et expliquer pourquoi certaines choses prennent du temps. C'est un exercice d'assertivité permanent."

Pour surmonter ces défis, tous les interviewés insistent sur l'importance de la collaboration, de la documentation et de l'automatisation des tâches répétitives. Le partage de connaissances au sein de l'équipe et la mise en place de bonnes pratiques collectives sont des leviers essentiels pour faire face à la complexité.

Ces défis me semblent stimulants intellectuellement. Mon profil me prédispose à bien gérer la complexité technique et à maintenir un haut niveau de qualité. En revanche, je dois progresser sur ma capacité à gérer les attentes des parties prenantes et à communiquer de manière plus affirmée lorsque c'est nécessaire. L'équilibre entre qualité et vitesse résonne particulièrement avec ma personnalité : j'ai naturellement tendance à privilégier la qualité, mais je dois apprendre à prioriser et à accepter parfois des solutions "suffisamment bonnes" plutôt que parfaites.

Question 6 : Quelles sont les évolutions technologiques majeures que vous avez observées ces dernières années et comment vous êtes-vous adapté ?

Mon collègue senior évoque la transition majeure vers le cloud : "Il y a 5 ans, tout était on-premise. Aujourd'hui, la majorité de nos infrastructures sont sur AWS. Cette migration a complètement changé notre façon de travailler : on ne gère plus de serveurs physiques, on utilise des services managés, on paie à l'usage. Ça demande un apprentissage constant."

L'architecte Data mentionne l'émergence des architectures data mesh et data fabric : "On est passé d'une approche centralisée avec un data warehouse unique à des architectures décentralisées où chaque domaine métier gère ses propres données. C'est un changement de paradigme complet qui nécessite de repenser la gouvernance et la qualité des données."

Mon ancienne camarade DataOps Engineer insiste sur la conteneurisation et Kubernetes : "Avant, on déployait tout manuellement. Maintenant, tout est conteneurisé, orchestré par Kubernetes. Ça a révolutionné notre façon de gérer les environnements et de scaler les applications."

La connaissance manager évoque l'explosion du temps réel : "Les métiers ne veulent plus attendre le lendemain matin pour avoir leurs indicateurs. Ils veulent du temps réel ou du quasi-temps réel. Ça nous a obligé à adopter des technologies de streaming comme Kafka, alors qu'avant on ne faisait que du batch."

Analyse personnelle : Ces témoignages me font prendre conscience de l'importance de la veille technologique et de l'apprentissage continu. Le métier évolue tellement vite qu'il faut constamment se former. Mon plan d'action doit donc inclure du temps dédié à l'expérimentation de nouvelles technologies, pas seulement pour suivre les modes, mais pour comprendre quels problèmes elles résolvent vraiment.

Question 7 : Quels conseils donneriez-vous à quelqu'un qui débute dans le métier de Data Engineer ?

Mon collègue senior insiste sur l'importance des fondamentaux : "Maîtrise vraiment bien SQL et Python avant de te lancer dans des technos plus complexes. Beaucoup de juniors veulent tout de suite faire du Spark ou du Kafka, mais sans bases solides, c'est compliqué. Et surtout, comprends comment fonctionnent les bases de données, la modélisation dimensionnelle, les index, les optimisations de requêtes."

L'architecte Data recommande de développer une culture de la qualité dès le début : "Prends l'habitude de tester ton code, de documenter ce que tu fais, de faire des revues de code. Ces bonnes pratiques te feront gagner énormément de temps par la suite et te rendront beaucoup plus efficace."

Mon ancienne camarade suggère de ne pas avoir peur de l'échec : "Tu vas passer des heures sur des bugs incompréhensibles, tu vas casser des trucs en production, c'est normal. L'important c'est d'apprendre de chaque erreur et de mettre en place des mécanismes pour ne pas les reproduire. Et surtout, demande de l'aide quand tu bloques, ne reste pas seul dans ton coin."

La connaissance manager conseille de s'intéresser au métier et pas seulement à la technique : "Va voir les utilisateurs finaux de tes dashboards, comprends ce qu'ils en font, demande-leur si ça répond à leurs besoins. Un Data Engineer qui comprend le business est beaucoup plus efficace qu'un pur technicien."

Ces conseils résonnent fortement avec mes axes de développement. L'importance accordée aux fondamentaux me rassure sur ma formation actuelle au CESI. Le conseil sur la culture de la qualité confirme que ma rigueur naturelle est un atout. En revanche, le point sur "ne pas rester seul dans son coin" me challenge directement sur ma tendance à l'isolement quand je bloque sur un problème.

Question 8 : Comment gérez-vous l'équilibre entre innovation technologique et maintien de la stabilité des systèmes en production ?

Mon collègue senior explique : "On ne peut pas tout refaire à chaque fois qu'une nouvelle techno sort. Il faut évaluer le rapport coût/bénéfice. Parfois, garder une solution qui fonctionne, même si elle n'est pas la plus à la mode, c'est le bon choix. L'innovation pour l'innovation, ça ne sert à rien."

L'architecte Data décrit une approche structurée : "On a mis en place un processus de POC (Proof of Concept) pour les nouvelles technologies. Avant d'adopter un nouvel outil, on le teste sur un petit périmètre, on évalue ses performances, sa facilité de maintenance, son coût. Ce n'est qu'après validation qu'on le généralise."

Mon ancienne camarade mentionne l'importance de la documentation et du transfert de compétences : "Quand on introduit une nouvelle techno, il faut former toute l'équipe. Si une seule personne la maîtrise, ça devient un point de risque. On organise des sessions de partage, on documente les choix qu'on a faits et pourquoi."

La connaissance manager ajoute une dimension organisationnelle : "Il faut aussi convaincre le management. L'innovation a un coût : temps de formation, risques de migration, période d'adaptation. Il faut savoir argumenter sur la valeur business que ça apporte, pas juste dire que c'est cool techniquement."

Analyse personnelle : Cette question m'interpelle car j'ai naturellement tendance à être attiré par les nouvelles technologies. Ces retours me rappellent l'importance du pragmatisme : il ne s'agit pas d'utiliser la dernière techno à la mode, mais de choisir les outils qui résolvent réellement les problèmes. Je dois développer cette capacité d'évaluation critique et ne pas me laisser séduire uniquement par l'aspect technique.

Question 9 : Quel est votre rapport au travail en équipe et comment collaborez-vous avec les autres profils data (Data Analysts, Data Scientists, etc.) ?

La collaboration apparaît comme un élément central du métier, bien plus que ce que j'imaginai initialement.

Mon collègue senior décrit des interactions quotidiennes : "Je travaille constamment avec les Data Analysts. Ils me remontent des incohérences dans les données, je les aide à optimiser leurs requêtes, on échange sur la meilleure façon de modéliser certaines informations. C'est un aller-retour permanent."

L'architecte Data évoque la collaboration avec les Data Scientists : "Les Data Scientists ont besoin de données propres et structurées pour entraîner leurs modèles. Mon rôle est de leur fournir des pipelines fiables qui leur donnent accès à ces données. Mais il faut aussi comprendre leurs contraintes : les volumes dont ils ont besoin, la fréquence de rafraîchissement, les transformations spécifiques."

Mon ancienne camarade insiste sur l'importance des rituels d'équipe : "On fait des stand-ups quotidiens pour se synchroniser, des revues de code systématiques, des

rétrospectives pour améliorer nos process. Ces rituels créent une cohésion d'équipe et permettent de partager les connaissances."

La connaissance manager souligne les défis de la collaboration : "Les tensions peuvent apparaître quand les priorités divergent. Les Data Scientists veulent aller vite et expérimenter, nous on veut de la stabilité et de la qualité. Il faut trouver des compromis, définir des règles claires, communiquer."

Ces témoignages confirment que le métier de Data Engineer est loin d'être solitaire. Cette réalité me challenge directement sur mon profil introverti et ma tendance à travailler seul. Je vais devoir sortir de ma zone de confort et développer mes compétences collaboratives. La mention des rituels d'équipe me rassure : avoir un cadre structuré pour les échanges me conviendra mieux que des interactions informelles permanentes.

Question 10 : Quelles sont les erreurs les plus fréquentes que vous voyez chez les Data Engineers juniors et comment les éviter ?

Mon collègue senior pointe le manque de tests : "Beaucoup de juniors développent leurs pipelines, vérifient que ça marche sur quelques exemples et déploient en production. Puis ils découvrent des cas limites qui font tout planter. Il faut toujours tester exhaustivement : données nulles, formats inattendus, volumes importants, etc."

L'architecte Data évoque le sur-engineering : "Les juniors ont tendance à vouloir construire des systèmes ultra-complexes et ultra-flexibles 'au cas où'. Mais la complexité est l'ennemi de la fiabilité. Il vaut mieux commencer simple et faire évoluer progressivement en fonction des besoins réels."

Mon ancienne camarade mentionne la négligence de la documentation : "On se dit toujours qu'on documentera plus tard, mais on ne le fait jamais. Résultat, 6 mois après, personne ne comprend pourquoi on a fait tel choix ou comment fonctionne tel pipeline. La documentation doit se faire en même temps que le développement."

La connaissance manager pointe le manque d'anticipation sur la scalabilité : "Un pipeline qui fonctionne bien avec 1000 lignes peut complètement s'effondrer avec 1 million de lignes. Il faut toujours se demander : qu'est-ce qui se passe si le volume est multiplié par 10 ? par 100 ? Et concevoir en conséquence."

Une autre erreur fréquente mentionnée par plusieurs interviewés est le manque de monitoring : "On développe un pipeline, on le met en production, et on ne surveille

rien. Puis un jour, il tombe en panne depuis 3 jours et personne ne s'en est rendu compte. Il faut toujours mettre en place des alertes."

Ces erreurs sont autant de pièges que je dois éviter. Le point sur le sur-engineering résonne particulièrement avec mon profil : j'ai tendance à vouloir faire les choses parfaitement dès le départ, quitte à créer de la complexité inutile. Je dois apprendre à itérer, à commencer simple et à complexifier seulement quand c'est nécessaire. Le point sur la documentation est également un rappel important : ma tendance à me concentrer sur le technique ne doit pas se faire au détriment de la transmission de connaissances.

Conclusion générale des entretiens

Ces dix questions ont permis de dresser un portrait complet et nuancé du métier de Data Engineer. Au-delà des compétences techniques, ce qui ressort fortement de ces échanges, c'est l'importance des dimensions humaines, organisationnelles et stratégiques du métier.

Le Data Engineer n'est pas un développeur isolé qui écrit du code dans son coin. C'est un facilitateur qui doit comprendre les besoins métier, collaborer avec de multiples parties prenantes, faire preuve de pragmatisme dans ses choix techniques, et maintenir un équilibre délicat entre innovation et stabilité.

Ces entretiens confirment que mon profil présente des forces réelles (rigueur, autonomie, capacité d'apprentissage) mais aussi des axes de développement clairs (communication proactive, collaboration). Les deux prochaines années de ma formation CESI seront l'occasion de travailler consciemment sur ces axes pour me préparer au mieux à l'exercice du métier.

3. ENQUÊTE SUR SOI - ANALYSE INTROSPECTIVE

3.1. Feedback 360° : Le regard de l'entourage

Pour compléter l'analyse du métier par une compréhension fine de mon profil personnel, j'ai sollicité les retours de mon entourage professionnel et personnel. Cette démarche de feedback 360° visait à identifier mes forces et mes axes de progression tels qu'ils sont perçus par les personnes qui me côtoient au quotidien. J'ai interrogé ma famille, mes collègues de travail chez Hérault Logement, mes camarades de promotion CESI et quelques amis proches.

Les retours récoltés dessinent un portrait cohérent de ma personnalité et de ma manière de travailler. Les qualificatifs qui reviennent le plus fréquemment sont : autonome, rigoureux, calme, réservé, fiable et méthodique.

L'autonomie est la caractéristique la plus souvent citée. Mes collègues chez Hérault Logement soulignent ma capacité à prendre en charge des sujets complexes sans avoir besoin d'un encadrement permanent. Cette autonomie se manifeste par ma capacité à définir moi-même mes priorités, à rechercher les informations nécessaires et à mener les projets jusqu'à leur terme. Mon maître d'apprentissage apprécie particulièrement cette qualité qui lui permet de me confier des missions avec confiance, sachant que je saurai gérer les difficultés de manière proactive.

La rigueur dans le travail est également largement reconnue. Mes camarades de CESI notent mon souci du détail et ma volonté de faire les choses correctement. Cette rigueur se traduit par une attention particulière portée à la qualité des livrables, à la documentation du travail effectué et au respect des standards techniques. Dans mes projets de développement, je prends systématiquement le temps de tester mes solutions, de vérifier les cas limites et de m'assurer que tout fonctionne avant de considérer une tâche comme terminée.

Le calme et la réserve sont des traits de personnalité que tous mes interlocuteurs relèvent. Ma famille décrit une personne posée qui ne se laisse pas facilement déstabiliser par les difficultés. Cette stabilité émotionnelle est perçue comme un atout dans la gestion du stress et des situations de crise. Mes collègues apprécient mon attitude mesurée dans les réunions et ma capacité à analyser les problèmes avec du recul avant de réagir.

Cependant, cette réserve peut parfois être perçue de manière ambivalente. Si elle est vue comme une forme de maturité et de sang-froid par certains, d'autres soulignent

qu'elle peut donner l'impression d'un manque de proactivité dans la communication. Plusieurs retours mentionnent que je pourrais davantage partager mes difficultés, mes interrogations ou mes succès. Cette discrétion, bien qu'elle ne soit pas un défaut en soi, peut créer des zones d'ombre dans la compréhension de mon travail et de mes besoins.

La fiabilité est un autre point fort régulièrement mentionné. Quand je m'engage sur une tâche ou un délai, mes interlocuteurs savent qu'ils peuvent compter sur moi. Cette fiabilité se construit sur ma capacité à évaluer correctement la charge de travail, à anticiper les difficultés et à communiquer suffisamment tôt si un problème risque d'impacter les délais. Cette qualité est particulièrement valorisée en entreprise où la coordination entre équipes nécessite que chacun tienne ses engagements.

Les axes de progression identifiés par mon entourage concernent principalement ma communication. Plusieurs personnes suggèrent que je pourrais davantage prendre la parole en réunion, partager mes idées plus spontanément et solliciter plus régulièrement l'avis des autres. Cette réserve communicationnelle, si elle ne pose pas de problème majeur actuellement, pourrait limiter mon impact dans des postes à responsabilité où la capacité à convaincre et à fédérer devient déterminante.

3.2. Profil de personnalité et soft skills

Pour approfondir la connaissance de mon profil, j'ai réalisé un test de personnalité basé sur le modèle Myers-Briggs (MBTI). Les résultats m'ont classé dans la catégorie ISTP-T (Virtuoso turbulent), avec les scores suivants :

- 68% Introverti : Je préfère les interactions sociales en petit comité et j'ai besoin de moments de solitude pour me ressourcer. Les échanges profonds et significatifs me conviennent mieux que les conversations superficielles ou les grands groupes.
- 53% Observant : J'ai une approche pragmatique et ancrée dans la réalité. Je privilégie les faits concrets et l'expérience pratique plutôt que les théories abstraites. Cette orientation me rend particulièrement à l'aise dans les environnements techniques où l'expérimentation et le concret dominent.
- 57% Thinking : Mes décisions sont principalement guidées par la logique et l'analyse objective plutôt que par les émotions. Cette rationalité me permet d'aborder les problèmes de manière structurée et de maintenir mon calme dans les situations stressantes.

- 76% Prospecting : Je privilégie la flexibilité et l'adaptabilité plutôt que la planification stricte. Cette tendance se manifeste par ma capacité à gérer plusieurs projets en parallèle et à m'adapter rapidement aux changements de contexte. Cependant, le test souligne que cette flexibilité peut parfois poser problème dans les environnements qui exigent une rigueur procédurale stricte.
- 56% Turbulent : Cette dimension indique une tendance à l'auto-critique et à la recherche constante d'amélioration. Je suis conscient de mes imperfections et je cherche activement à progresser, ce qui peut parfois se traduire par un manque de confiance en moi dans certaines situations.

L'interprétation de ce profil révèle plusieurs traits pertinents pour le métier de Data Engineer. Mon introversion et mon orientation vers la réflexion correspondent bien à un métier qui nécessite concentration et analyse approfondie. Ma préférence pour le concret et l'expérimentation fait de moi quelqu'un qui apprend efficacement par la pratique, ce qui est un atout dans un domaine technique en constante évolution.

La dimension "Thinking" confirme mon approche rationnelle et méthodique du travail. Cette logique dominante me permet d'analyser les problèmes de manière objective et de construire des solutions robustes. Cependant, le test souligne que cette rationalité doit être équilibrée par une attention aux dimensions humaines et relationnelles, notamment dans les interactions avec les équipes métier qui ne partagent pas nécessairement cette approche analytique.

La composante "Prospecting" est particulièrement intéressante car elle révèle une tension interne. D'un côté, j'ai une forte capacité d'adaptation et j'apprécie la flexibilité. De l'autre, le métier de Data Engineer exige souvent une rigueur et une structuration importantes, notamment dans la documentation et les processus. Cette tension explique peut-être certaines difficultés que je peux rencontrer dans les phases de documentation ou de formalisation, tâches que je trouve moins stimulantes que la résolution de problèmes techniques.

La dimension "Turbulent" indique une tendance à l'auto-critique qui, si elle est bien gérée, peut être un moteur de progression. Cette conscience de mes limites me pousse à apprendre continuellement et à chercher à m'améliorer. Cependant, elle peut aussi se manifester par un manque de confiance en moi, notamment dans ma communication avec les autres. Le test suggère que je dois travailler sur ma capacité à reconnaître mes succès et à affirmer mes positions avec plus d'assurance.

3.3. Système de valeurs et motivations profondes

Pour compléter l'analyse de mon profil, il est important d'identifier mes valeurs profondes et mes motivations au travail. Ces éléments constituent le socle de mon engagement professionnel et orienteront mes choix de carrière à long terme.

L'impact est ma valeur centrale. Ce qui me motive fondamentalement dans mon travail, c'est la conviction que ce que je fais a un effet positif sur les personnes et sur l'organisation. Dans mon expérience actuelle chez Hérault Logement, cette motivation se concrétise par le développement de dashboards Power BI qui permettent aux équipes de prendre de meilleures décisions, d'optimiser leurs ressources et d'améliorer la qualité de service. Savoir que mon travail facilite le quotidien des collègues et contribue à l'efficacité de l'entreprise est un moteur puissant pour moi.

Cette valeur d'impact rejoint directement le métier de Data Engineer. En construisant des pipelines de données fiables et performants, en garantissant la qualité de l'information et en facilitant l'accès aux données pour les équipes d'analyse, le Data Engineer a un effet multiplicateur sur la capacité de l'organisation à prendre des décisions éclairées. Cette dimension me motive particulièrement.

L'autonomie constitue une autre valeur forte pour moi. J'apprécie de pouvoir organiser mon travail comme je l'entends, de choisir mes méthodes et mes outils, et de progresser à mon rythme. Cette autonomie ne signifie pas isolement : je valorise la collaboration et les échanges avec mes collègues, mais j'ai besoin de liberté dans l'exécution de mes tâches. Le métier de Data Engineer offre généralement cette autonomie, notamment dans les phases de conception et de développement où chaque ingénieur peut exprimer sa créativité technique.

La technicité est également une source de motivation importante. J'apprécie les défis intellectuels, la résolution de problèmes complexes et l'apprentissage continu. Le domaine de la data évolue rapidement, avec l'émergence régulière de nouvelles technologies et de nouveaux paradigmes. Cette dynamique correspond parfaitement à mon besoin de stimulation intellectuelle. La perspective de devoir apprendre en permanence, loin de me rebuter, me stimule.

La qualité du travail est une valeur que je place très haut. Je ne me satisfais pas de solutions approximatives ou de livrables médiocres. Cette exigence personnelle se traduit par le temps que je consacre à vérifier, tester et améliorer mes réalisations. Si cette quête

de qualité est un atout dans un métier où les erreurs peuvent avoir des conséquences importantes, elle doit être équilibrée par une capacité à prioriser et à accepter parfois des compromis.

La sobriété et l'efficacité constituent une valeur émergente pour moi. À mesure que je prends conscience des enjeux environnementaux, je suis de plus en plus sensible à la nécessité d'optimiser les ressources et de limiter le gaspillage. Dans le domaine de la data, cette valeur se traduit par une attention portée à l'optimisation des requêtes, au juste dimensionnement des infrastructures et à la mise en place de politiques de rétention raisonnables. Cette préoccupation rejoint les enjeux de sobriété numérique identifiés dans l'analyse du métier.

Le respect et la bienveillance dans les relations professionnelles sont également importants pour moi. J'apprécie de travailler dans des environnements où les personnes se traitent avec respect mutuel, où l'erreur est perçue comme une opportunité d'apprentissage plutôt que comme une faute, et où la collaboration prime sur la compétition. Cette valeur influence mes choix dans la sélection de mes futurs employeurs.

4. LIEN ENTRE ENQUÊTE METIER ET ENQUÊTE SUR SOI

Pour combler les écarts identifiés durant mes années A4 et A5, ma stratégie repose sur un développement équilibré entre montée en compétences techniques et renforcement des soft skills, deux dimensions également critiques pour réussir dans le métier de Data Engineer.

Sur le plan technique, je souhaite construire une expertise progressive sur les technologies data modernes tout en conservant une vision d'ensemble de la chaîne de valeur. Cela passera par une montée en niveau sur les technologies cloud* (en visant une certification AWS ou Azure), l'apprentissage des outils de traitement distribué comme Spark, et la découverte des solutions d'orchestration de pipelines telles qu'Airflow. Mon objectif n'est pas de tout maîtriser immédiatement, mais d'acquérir une compréhension suffisamment solide pour être rapidement opérationnel en sortie d'école. Je privilégierai l'apprentissage par la pratique en développant des projets personnels concrets qui me permettront d'expérimenter ces technologies dans un environnement sans pression.

Parallèlement, je souhaite approfondir mes connaissances sur les enjeux de sobriété numérique* et d'optimisation des ressources, dimension qui correspond à mes valeurs personnelles et qui représente un axe de différenciation pour la suite de ma carrière. Cela passera par une veille active sur les bonnes pratiques de Green IT* et par l'application concrète de ces principes dans mes projets chez Hérault Logement, notamment en analysant et optimisant la consommation de ressources de mes dashboards Power BI.

Concernant mon développement personnel, mon plan d'action se concentre sur deux axes prioritaires : la communication proactive et la gestion de l'équilibre entre qualité et pragmatisme. Je m'engage à structurer davantage mes comptes-rendus d'activité et à partager plus régulièrement l'avancement de mes travaux avec mon maître d'apprentissage et mes collègues, même lorsque cela ne m'est pas explicitement demandé. L'objectif est de transformer ma réserve naturelle en une communication professionnelle efficace, sans pour autant renier ma personnalité introvertie.

Je vais également travailler sur ma capacité à prendre la parole lors des réunions techniques, en m'imposant l'objectif de formuler au moins une question ou un commentaire par réunion. Cette pratique régulière me permettra progressivement de gagner en assurance et de dépasser mon appréhension à m'exprimer en groupe. De même, je solliciterai plus facilement de l'aide lorsque je bloque sur un problème technique, plutôt que de rester seul dans mon coin pendant des heures.

4. CIBLE MÉTIER

À l'issue de ma formation d'ingénieur en 2028, mon objectif est de débiter ma carrière en tant que Data Engineer au sein d'une entreprise où la donnée occupe une place stratégique. Ce métier représente pour moi l'équilibre idéal entre expertise technique, impact organisationnel et évolution continue.

Définition du poste et missions principales

Le Data Engineer conçoit, développe et maintient les infrastructures qui permettent la collecte, le stockage, le traitement et la mise à disposition des données au sein d'une organisation. Concrètement, mon rôle consistera à orchestrer le "voyage de la donnée" depuis sa création jusqu'à son exploitation par les équipes d'analyse ou les modèles d'intelligence artificielle.

Plusieurs facteurs motivent mon choix de ce métier. D'abord, la diversité des missions. Le Data Engineer ne se cantonne pas à une tâche répétitive : il touche à l'infrastructure, au développement, à l'architecture et au conseil. Cette variété correspond parfaitement à mon besoin de stimulation intellectuelle et à mon goût pour l'apprentissage continu.

Ensuite, l'impact organisationnel. En garantissant la disponibilité et la qualité des données, le Data Engineer crée les conditions de possibilité pour que l'ensemble de l'organisation puisse prendre des décisions éclairées. Cet impact indirect mais fondamental résonne avec ma valeur de contribution positive.

La dimension technique du métier me séduit également. Le data engineering mobilise des technologies variées et en constante évolution. Cette complexité technique représente un défi permanent qui m'empêchera de m'ennuyer et me poussera à progresser continuellement.

Enfin, les perspectives d'évolution sont excellentes. Un Data Engineer peut évoluer vers des postes d'architecte data, de lead data engineer, de data platform engineer ou même de CTO. Cette diversité de trajectoires possibles m'offre de la flexibilité pour construire ma carrière en fonction de mes aspirations futures.

Lien avec les Objectifs de Développement Durable

Le métier de Data Engineer s'inscrit naturellement dans plusieurs Objectifs de Développement Durable définis par l'ONU.

ODD 9 : Industrie, innovation et infrastructure - En construisant des infrastructures data robustes et évolutives, le Data Engineer favorise l'innovation au sein des organisations. Des données de qualité, accessibles et bien gouvernées constituent le socle de l'innovation data-driven qui permet aux entreprises de développer de nouveaux produits et services.

ODD 12 : Consommation et production responsables - C'est l'ODD qui résonne le plus avec mes préoccupations personnelles. En tant que Data Engineer, j'ai un pouvoir direct sur l'optimisation des ressources informatiques. Chaque choix d'architecture, chaque optimisation de requête, chaque politique de rétention des données a un impact sur la consommation énergétique et l'empreinte carbone de l'organisation.

ODD 8 : Travail décent et croissance économique - En améliorant l'efficacité opérationnelle des organisations grâce à la donnée, le Data Engineer contribue indirectement à leur performance économique. Des décisions mieux informées permettent d'optimiser les processus, de réduire les gaspillages et d'identifier de nouvelles opportunités de croissance.

Vision à 5-10 ans

À court terme (2-3 ans), mon objectif est de consolider mon expertise technique sur l'ensemble de la stack data engineering moderne et de développer ma capacité à travailler en équipe sur des projets complexes.

À moyen terme (5 ans), je souhaite évoluer vers un poste de Data Engineer senior où je pourrai prendre en charge des sujets d'architecture et commencer à encadrer des profils juniors. Cette évolution me permettra de développer mes compétences en leadership technique.

À plus long terme (10 ans), plusieurs trajectoires m'attirent : devenir architecte data pour avoir une vision globale des systèmes d'information, évoluer vers un poste de lead data engineer pour manager une équipe, ou me spécialiser dans un domaine de niche comme le streaming en temps réel ou l'optimisation des performances.

Quelle que soit l'évolution choisie, mon fil conducteur restera le même : utiliser mes compétences techniques pour créer de la valeur, avoir un impact positif sur les organisations et contribuer à un numérique plus sobre et plus responsable.

5. BIBLIOGRAPHIE ET SOURCES

6.1. Ressources documentaires et bibliographiques

Études sur le marché de l'emploi :

APEC (Association pour l'Emploi des Cadres) : *Fiche métier : Data Engineer*. Analyse des compétences et du marché de l'emploi.

Lien : <https://www.apec.fr/tous-nos-metiers/informatique/data-engineer.html>

IESF (Ingénieurs et Scientifiques de France) : *Enquête nationale 2024 sur la situation des ingénieurs*. Statistiques de salaires et de responsabilités.

Lien : https://www.iesf.fr/752_p_50331/une-enquete-de-reference.html

France Travail : *Étude métiers : ingénieur data*. Perspectives d'emploi et compétences recherchées.

Lien : <https://www.francetravail.fr/>

Technologies et pratiques du data engineering :

Databricks : *The Big Book of Data Engineering*. Guide complet sur les architectures data modernes.

Lien : <https://www.databricks.com/resources/ebook/big-book-of-data-engineering>

Google Cloud : *Data Engineering on Google Cloud Platform*. Documentation sur les bonnes pratiques cloud.

Lien : <https://cloud.google.com/architecture/data-engineering>

Apache Software Foundation : *Documentation Apache Spark, Kafka et Airflow*. Références techniques des outils majeurs du data engineering.

Lien : <https://spark.apache.org/>, <https://kafka.apache.org/>, <https://airflow.apache.org/>

Sobriété numérique et développement durable :

The Shift Project : *Déployer la sobriété numérique*. Rapport sur l'impact environnemental du numérique.

Lien : <https://theshiftproject.org/publications/deployer-sobriete-numerique/>

GreenIT.fr : *L'empreinte environnementale du numérique mondial*. Étude technique sur l'éco-conception.

Lien : <https://www.greenit.fr/etude-empreinte-environnementale-du-numerique-mondial/>

ADEME : *La face cachée du numérique*. Guide pratique pour réduire l'impact environnemental.

Lien : <https://www.ademe.fr/>

Gouvernance des données et réglementation :

CNIL (Commission Nationale de l'Informatique et des Libertés) : *Guide RGPD pour les développeurs*.

Lien : <https://www.cnil.fr/fr/guide-rgpd-du-developpeur>

DAMA International : *DMBOK (Data Management Body of Knowledge)*. Référentiel des bonnes pratiques de gestion des données.

Objectifs de Développement Durable :

ONU : *Les 17 Objectifs de Développement Durable - Agenda 2030*.

Lien : <https://www.un.org/sustainabledevelopment/fr/objectifs-de-developpement-durable/>

Tests de personnalité :

16Personalities : *Test de personnalité MBTI*.

Lien : <https://www.16personalities.com/fr>

Glossaire

Agile : Méthode de gestion de projet itérative privilégiant la flexibilité, la collaboration et la livraison fréquente de fonctionnalités.

Airflow : Outil open-source d'orchestration de workflows permettant de programmer et de monitorer l'exécution de pipelines de données complexes.

Alerting : Système de notification automatique qui prévient les équipes techniques lorsqu'un seuil ou une condition anormale est détecté dans un système.

Ansible : Outil d'automatisation IT permettant de configurer des systèmes et de déployer des applications de manière reproductible.

Auto-ML (Automated Machine Learning) : Technologies qui automatisent le processus de sélection, d'entraînement et d'optimisation de modèles de machine learning.

Auto-scaling : Mécanisme qui ajuste automatiquement les ressources informatiques (serveurs, puissance de calcul) en fonction de la charge.

AWS (Amazon Web Services) : Plateforme cloud leader proposant des services d'infrastructure, de stockage et de traitement de données à la demande.

Batch : Mode de traitement où les données sont collectées et traitées par lots à intervalles réguliers, par opposition au streaming.

Big Data : Technologies et méthodes permettant le stockage, le traitement et l'analyse de volumes de données massifs dépassant les capacités des outils traditionnels.

CI/CD (Continuous Integration / Continuous Deployment) : Pratiques DevOps consistant à automatiser l'intégration du code et son déploiement en production de manière continue et fiable.

Cloud : Modèle d'infrastructure IT où les ressources (serveurs, stockage, applications) sont accessibles à la demande via internet plutôt que sur site.

Cloud-native : Architecture logicielle conçue spécifiquement pour tirer parti des avantages du cloud (scalabilité, résilience, élasticité).

Cloud provider : Fournisseur de services informatiques via internet (stockage, calcul, bases de données). Les principaux sont AWS, Microsoft Azure et Google Cloud Platform.

Conteneurisation : Technologie (comme Docker) permettant d'encapsuler une application et ses dépendances dans un "conteneur" portable qui peut s'exécuter de manière identique sur n'importe quel environnement.

Data fabric : Architecture qui crée une couche d'abstraction au-dessus des données distribuées dans différents systèmes, permettant un accès unifié sans nécessiter de centralisation physique.

Data governance : Ensemble de processus, de rôles et de règles garantissant la qualité, la sécurité et la conformité des données au sein d'une organisation.

Data lake : Entrepôt de stockage centralisé permettant de stocker des données structurées et non structurées à grande échelle dans leur format brut.

Data lineage : Traçabilité du cycle de vie des données permettant de comprendre d'où viennent les données, comment elles ont été transformées et où elles sont utilisées.

Data mesh : Approche architecturale décentralisée où chaque domaine métier est responsable de ses propres données, plutôt que de tout centraliser dans une plateforme unique.

Data quality : Ensemble de pratiques visant à garantir que les données sont exactes, complètes, cohérentes et fiables pour leur usage prévu.

Data self-service : Environnement permettant aux utilisateurs métier non techniques d'accéder de manière autonome aux données et de créer leurs propres analyses sans passer systématiquement par l'IT.

Data warehouse : Base de données centralisée optimisée pour l'analyse, regroupant des données provenant de multiples sources et organisées selon un modèle structuré.

DataOps : Application des principes DevOps au domaine de la donnée pour améliorer la qualité, réduire les temps de cycle et faciliter la collaboration entre équipes.

DevOps : Culture et ensemble de pratiques visant à unifier le développement logiciel et l'administration des infrastructures IT pour accélérer les livraisons.

Dimensionnel (modèle) : Approche de modélisation de données pour l'analyse, organisant l'information en tables de faits (métriques) et tables de dimensions (contexte).

Docker : Technologie de conteneurisation permettant d'empaqueter une application et ses dépendances dans un conteneur portable et reproductible.

ESN (Entreprise de Services du Numérique) : Société proposant des services informatiques et de conseil, anciennement appelée SSII.

ETL (Extract, Transform, Load) : Processus informatique consistant à extraire des données de sources diverses, les transformer selon des règles métier et les charger dans un système cible.

Événementielle (architecture) : Architecture logicielle basée sur la production, la détection et la réaction à des événements en temps réel.

Gap Analysis : Méthode d'évaluation comparant l'état actuel d'une situation (compétences, performances) à l'état souhaité pour identifier les écarts à combler.

GLPI : Logiciel libre de gestion de parc informatique et de help desk permettant l'inventaire et le suivi des incidents IT.

Green IT : Ensemble de techniques visant à réduire l'empreinte écologique des technologies de l'information (optimisation du code, réduction de la consommation énergétique).

IaC (Infrastructure as Code) : Pratique consistant à gérer et provisionner l'infrastructure IT via du code plutôt que par des processus manuels.

Kafka : Plateforme de streaming distribuée permettant de publier, stocker et traiter des flux de données en temps réel à grande échelle.

Kubernetes (K8s) : Système open-source d'orchestration de conteneurs automatisant le déploiement, la mise à l'échelle et la gestion d'applications conteneurisées.

Legacy : Se dit de systèmes, applications ou technologies anciennes toujours en production mais dont les technologies sous-jacentes sont obsolètes ou dépassées.

MBTI (Myers-Briggs Type Indicator) : Test de personnalité identifiant 16 types psychologiques basés sur quatre dimensions : Extraversion/Introversion, Sensation/Intuition, Pensée/Sentiment, Jugement/Perception.

Monitoring : Surveillance continue des performances et de l'état de santé des systèmes IT pour détecter proactivement les anomalies et les incidents.

NoSQL : Catégorie de bases de données non relationnelles conçues pour gérer des données non structurées ou semi-structurées à grande échelle (MongoDB, Cassandra, Elasticsearch).

ODD (Objectifs de Développement Durable) : 17 objectifs définis par l'ONU dans l'Agenda 2030 pour répondre aux défis mondiaux (pauvreté, climat, inégalités, etc.).

On-premise : Se dit d'une infrastructure informatique hébergée et gérée localement par l'entreprise dans ses propres locaux, par opposition au cloud.

Open source : Logiciels dont le code source est public et peut être utilisé, modifié et redistribué librement, souvent développés de manière collaborative.

Orchestration : Coordination automatisée de tâches informatiques complexes et interdépendantes, comme l'exécution séquentielle ou parallèle de pipelines de données.

Pattern : Modèle de conception réutilisable qui fournit une solution éprouvée à un problème récurrent dans le développement logiciel ou l'architecture.

Pipeline : Séquence automatisée d'étapes de traitement de données, depuis leur collecte jusqu'à leur mise à disposition pour l'analyse ou l'exploitation.

Power BI : Outil de visualisation et de business intelligence de Microsoft permettant de créer des tableaux de bord interactifs et des rapports analytiques.

Proactivité : Capacité à anticiper les problèmes ou les besoins avant qu'ils ne surviennent, plutôt que de simplement réagir aux événements.

Promptant : Action de formuler des instructions textuelles à une intelligence artificielle générative pour obtenir un résultat spécifique.

Reverse engineering : Processus d'analyse d'un système pour en comprendre le fonctionnement interne, généralement lorsque la documentation est absente ou obsolète.

RGPD (Règlement Général sur la Protection des Données) : Réglementation européenne encadrant la collecte et le traitement des données personnelles pour protéger la vie privée des citoyens.

Roadmap : Planification stratégique présentant les grandes étapes et les jalons d'un projet ou d'un produit sur une période donnée.

Root cause : Cause fondamentale et originelle d'un problème, dont la correction empêche la réapparition de l'incident.

Scalabilité : Capacité d'un système à maintenir ses performances lorsque la charge ou le volume de données augmente significativement.

Scale-up : Startup en phase de croissance accélérée qui a dépassé le stade de l'amorçage et commence à se structurer.

Side project : Projet personnel développé en dehors du cadre professionnel, généralement pour expérimenter de nouvelles technologies ou approfondir des compétences.

Sobriété numérique : Démarche visant à réduire l'empreinte écologique et sociale du numérique en limitant la consommation de ressources (énergie, matériaux, bande passante).

Spark : Moteur de traitement de données distribué open-source permettant d'exécuter des calculs à grande échelle sur des clusters de machines.

Stack technique : Ensemble cohérent de technologies (langages, frameworks, outils) utilisées ensemble pour développer une application ou un système.

Streaming : Mode de traitement où les données sont traitées en continu au fur et à mesure de leur arrivée, permettant des analyses en temps réel.

Systémique (approche) : Méthode d'analyse considérant un système comme un ensemble d'éléments en interaction, plutôt que d'étudier chaque composant isolément.

Terraform : Outil d'Infrastructure as Code permettant de définir et de provisionner l'infrastructure cloud via des fichiers de configuration déclaratifs.