

Transition énergétique et technologies numériques : comment la donnée est utilisée pour la stratégie multi-énergies de TotalEnergies

Michel Lutz est depuis 2016 Chief Data Officer¹ pour la compagnie TotalEnergies. Son activité consiste ainsi à savoir quoi et comment faire pour mieux utiliser les données au niveau stratégique pour la compagnie. Il est aussi Head of Data au sein de TotalEnergies Digital Factory², filiale détenue à 100 % par TotalEnergies, et qui développe des solutions numériques pour la Compagnie. À ce titre, il pilote une équipe d'une quarantaine de data scientists, data engineers et autres experts en matière de data science, recherche opérationnelle, MLOps, IA générative, etc.

1. Responsable des données.

2. Usine digitale.

Introduction : la stratégie multi-énergies de TotalEnergies

TotalEnergies considère l'énergie comme l'un des enjeux majeurs de la société d'aujourd'hui et de celle de demain. À cet égard, la compagnie engage ce qu'elle considère être une importante transformation. Celle-ci repose sur deux grands axes pour répondre aux attentes des populations en termes de transition énergétique (*Figure 1*).

D'une part, nous travaillons sur la réduction et la compensation des émissions de dioxyde de carbone liées aux activités industrielles de la compagnie. Cela concerne la production et la distribution d'hydrocarbures, mais aussi toutes les nouvelles énergies. D'autre part, TotalEnergies accompagne des recherches et des projets dans le domaine des nouvelles énergies : l'électricité, le renouvelable, l'éolien, le solaire, l'hydrogène et la biomasse. La

chimie joue un grand rôle dans tous ces domaines dans lesquels existent énormément de données. L'utilisation conjointe des données et des technologies digitales est un levier important pour accompagner cette transformation.

La variété des données à notre disposition pour justement accompagner ces transformations de la compagnie est représentée sur la *Figure 2*.

Notre patrimoine dispose d'une grande variété de données avec des enjeux techniques très différents. **Notre cœur historique sur les données de sous-sol** représente des volumes énormes : nous disposons de 50 pétaoctets³ d'archives sismiques pouvant être traitées avec notre super ordinateur. Ces données représentent des grands enjeux pour la connaissance du sous-sol, et peuvent par

3. Unité de mesure de la mémoire, correspondant à un million de gigaoctets.

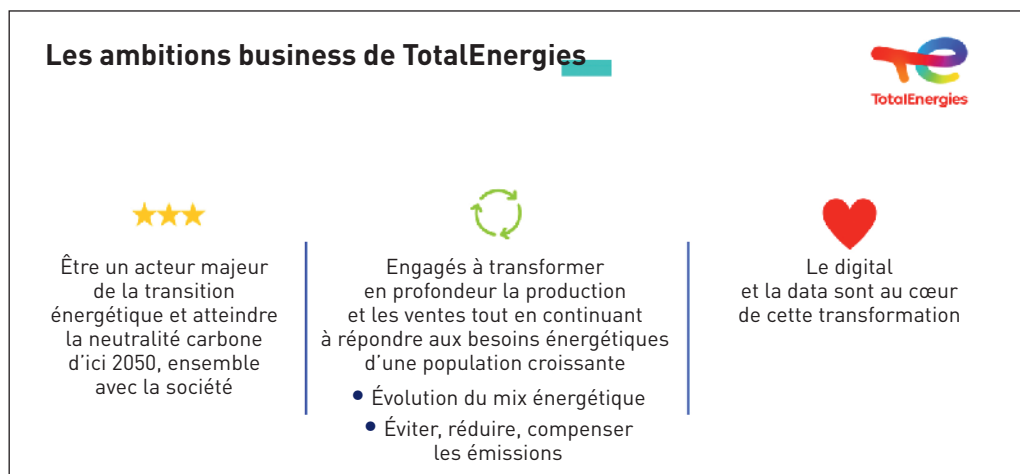


Figure 1

TotalEnergies et la transition énergétique.

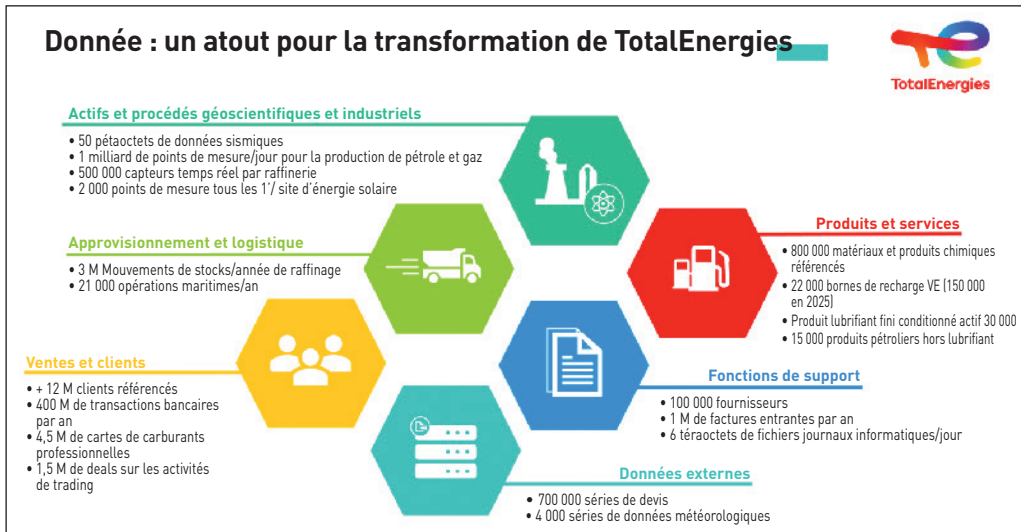


Figure 2

Patrimoine de données de TotalEnergies.

exemple être utilisées pour les projets de capture et de stockage de CO₂.

En ce qui concerne **notre activité industrielle**, nous avons des historiques de données importants sur les activités **de raffinage et de production de pétrole**. Nous travaillons aussi de plus en plus autour de ce qui concerne la production d'énergies renouvelables, notamment le solaire et l'éolien. Ce sont des unités industrielles instrumentées avec des dizaines de millions de capteurs et déployées dans le monde entier qui nous remontent de l'information sur le fonctionnement de nos usines de production d'énergies. C'est précieux, pour optimiser au mieux le fonctionnement de nos usines et diminuer les émissions de dioxyde de carbone.

Nous disposons aussi de données très variées parmi nos

12 millions de clients référencés dans les bases, par exemple pour tout ce qui est logistique client. C'est très important pour optimiser nos mécaniques de distribution d'énergie. Ce domaine est en pleine évolution avec le déploiement de bornes de rechargement.

1 L'utilisation des données et l'écosystème scientifique de la société

Les équipes travaillent de différentes façons sur l'utilisation de ces données. L'écosystème scientifique de la compagnie est représenté sur la **Figure 3**.

1.1. La recherche et le développement

Les équipes de R&D travaillent en amont. Ces équipes recherchent les solutions pour

les défis actuels et préparent les problématiques de demain. Certaines équipes maîtrisent et utilisent beaucoup l'intelligence artificielle et le *machine learning* dans les procédés chimiques, c'est le cas pour la génération de matériaux. Ce sont des domaines dans lesquels nous travaillons en partenariat avec des laboratoires universitaires.

1.2. Les équipes spécialisées

Nous avons des équipes spécialisées qui utilisent les nouvelles technologies ou les technologies « numériques » sur des problématiques métiers précises.

Historiquement, nous utilisons depuis longtemps la simulation numérique pour la modélisation du sous-sol et des réservoirs. Cependant, nous utilisons désormais de plus en plus les méthodes de *machine*

learning d'analyses de données pour enrichir les points de vue et les façons de traiter les problèmes.

Nous avons aussi des équipes historiques pour tout ce qui concerne la modélisation et l'optimisation des processus de raffinage.

Il y a enfin des équipes très spécialisées sur des sujets comme : la chimiométrie, la maîtrise statistique des processus dans les usines, la production d'énergies renouvelables, etc.

1.3. L'équipe Machine Learning (ML)

L'équipe data de la Digital Factory, que gère Michel Lutz, est spécialisée en *Learning engineering*⁴. Dans ce domaine, une compétence spécifique

4. Ingénierie de l'apprentissage machine.

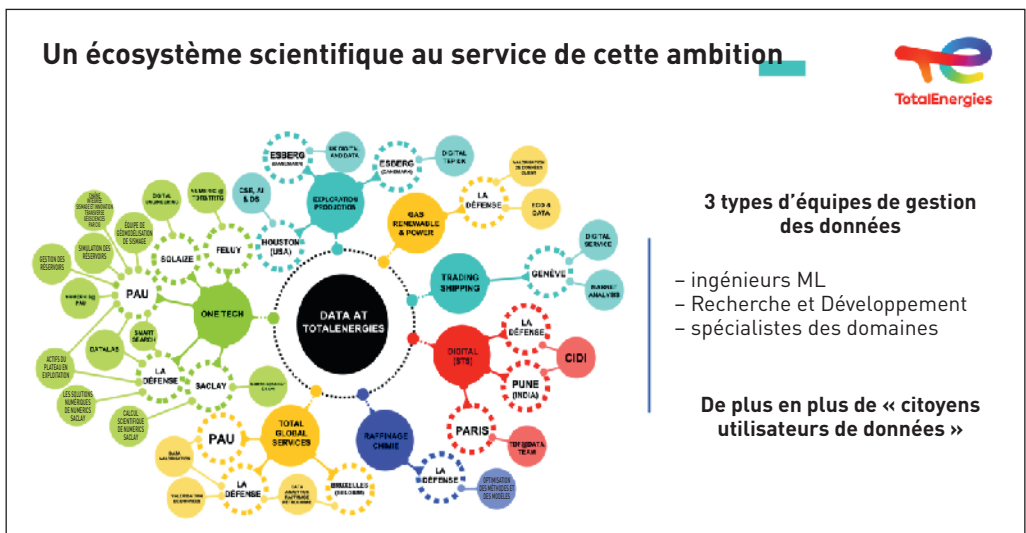


Figure 3

Un écosystème scientifique chez TotalEnergies.

est la mise en production des modèles dans des applications mises dans les mains de l'utilisateur. Nous devons élaborer des applications qui embarquent l'intelligence artificielle avec des interfaces d'utilisation développées avec des solutions logicielles robustes.

Par ailleurs, l'usage des données se démocratise de plus en plus : on appelle souvent cela *Citizen data* ou *Citizen data scientist*⁵.

Dans une société comme TotalEnergies, beaucoup de personnes, dans les métiers et dans les usines, ont de bonnes compétences scientifiques et

5. Donnée accessible. Ces professionnels utilisent régulièrement l'intelligence artificielle dans leur travail, sans pour autant l'avoir mise au point ni en étant expert.

savent manipuler les données. L'enjeu est donc de les rendre plus efficaces, de leur faciliter l'accès à la donnée pour qu'ils soient plus autonomes dans leur capacité à mieux décider au quotidien avec les bons outils et les bonnes données.

1.4. L'exemple du déploiement des panneaux solaires (Figure 4) : le programme SOLEAD

C'est l'équipe R&D qui à la base a travaillé sur la recherche d'algorithme de traitement d'images pour identifier à partir d'images satellites caractéristiques des zones potentiellement intéressantes pour développer les installations des panneaux solaires en fonction de différents facteurs dont le potentiel solaire de la zone.



Figure 4

L'installation de panneaux solaires : un exemple de coopération pour la recherche d'une zone potentiellement intéressante.

L'équipe (ML) travaille actuellement avec eux sur le résultat de cette recherche pour l'intégrer dans une application capable de délivrer l'information en temps réel, qui sera mise à la disposition des personnes de TotalEnergies qui développent l'activité solaire.

Nous développons en plus les algorithmes pour aider à la prise des décisions. Il existe tout un ensemble de données sur le contexte économique d'une zone géographique qui peuvent être intégrées, à partir desquelles on modélise les préférences des décideurs pour savoir si c'est intéressant ou non de mettre en place des panneaux solaires.

Cet ensemble permet de faire un produit qui sera ensuite utilisé de façon opérationnelle au quotidien pour la stratégie du déploiement solaire de TotalEnergies.

2 La Digital Factory Data Team (l'usine numérique)

L'équipe Data de la Digital Factory rassemble une quarantaine de personnes expertes en science des données appliquées et en ingénierie du *machine learning* pour trouver des solutions opérationnelles déployables et utiles pour les gens au quotidien.

Les compétences scientifiques doivent être pluridisciplinaires : modèles statistiques, *machine learning* (apprentissage automatique), recherche opérationnelle, compétences en génie logiciel et gestion du cycle de vie des modèles.

Il faut avoir une approche multi-méthodes pour trouver de

bonnes solutions mais ensuite, pour que ces solutions soient utiles, la partie développement de logiciels est importante.

Une dimension importante de notre travail est donc le développement de logiciel pour trouver des solutions répondant aux besoins des utilisateurs. Je citerai deux exemples de produits que nous avons réalisés.

Le premier est un logiciel de vision artificielle qui permet de reconnaître des gravures sur des tuyaux utilisés en industrie. Ces gravures étaient très difficiles à lire, donc nous avons créé un algorithme de vision artificielle que nous avons déployé sur téléphone portable afin de pouvoir scanner les tuyaux avec le téléphone portable. Après avoir eu la référence du tuyau, on peut ainsi l'utiliser pour optimiser les mouvements de déplacements de stocks.

L'autre exemple a été créé pour Saft, une filiale de TotalEnergies qui fabrique des batteries industrielles. Nous avons élaboré un algorithme qui fait de la prédiction de risques de non-qualité en sortie de chaîne de fabrication. Cet algorithme est basé sur toutes les données de capteurs de l'usine, et nous l'avons ensuite déployé sur un multimètre. Quand le technicien qualité fait une mesure avec le multimètre sur une batterie, cela appelle l'algorithme qui fait une prédiction de risque de non-qualité qui s'affiche dans l'usine sur un écran, et tous les techniciens qualité peuvent ainsi optimiser leurs processus d'inspection en fonction de

cette aide à la décision fournie par l'algorithme.

Ce type de travail nécessite une expertise assez spécifique en développement logiciel pour être capable d'amener les solutions jusqu'au bout dans les mains des utilisateurs.

Le deuxième axe de notre métier, au-delà de la partie technique, est de faire de l'intelligence artificielle pour et avec les utilisateurs : MLOps⁶.

Très souvent quand on fait un algorithme, il a au début un certain niveau de performance. Ensuite nous le mettons en production dans un logiciel, mais il va peut-être se tromper ou être confronté à des situations qu'il n'a jamais vues et pour lesquelles il pourra ne pas prendre une bonne décision. Nous devons construire des interfaces qui permettent justement de récupérer une

boucle de rétroaction avec l'utilisateur. Cela signifie que si l'algorithme n'a pas le comportement attendu, on aura un retour utilisateur pour avoir un entraînement continu, ce qui nécessite de mettre en place des mécaniques logicielles de réapprentissage assez particulières. C'est un objectif au cœur de ce qu'on appelle le MLOps.

Le troisième axe est un peu plus méthodologique mais néanmoins très important. Nous sommes très engagés dans ce qu'on appelle l'intelligence artificielle de confiance pour ne pas mettre en place des systèmes non maîtrisés qui pourraient prendre de façon autonome des mauvaises décisions, afin que nos utilisateurs aient confiance dans ce que nous faisons.

Pour réaliser cela, tout un cadre, qui s'inspire des bonnes pratiques qui existent sur le marché, a été mis en place et nous avons aussi des projets de régulation en cours pour s'assurer qu'on fait bien les choses et aussi donner un cadre de confiance entre les utilisateurs.

6. MLOps (*Machine Learning Operations*). Le MLOps peut se définir comme un ensemble de pratiques combinant le *machine learning*, le DevOps et le *Data Engineering*, qui vise à déployer et à maintenir les systèmes ML en production de manière fiable et efficace.

Conclusion

Perspectives stratégiques

En entreprise, les enjeux de la transformation data ne sont pas que de la science et pas que de la data science. Si on veut que ce soit bien fait, à l'échelle, et que ce soit impactant, il y a tout un ensemble de facilitateurs à mettre en place, et le rôle du Chief Data Officer se résume en 3 axes (*Figure 5*).

Technologies

Historiquement, les données proviennent des serveurs installés dans les usines et les systèmes opérationnels sont des données difficilement accessibles. Il y a donc toute une transformation technologique à mettre en place pour accéder plus facilement à ces données, les rendre plus facilement disponibles, pour enfin être capable de les utiliser avec les technologies qui permettent de faire de l'intelligence artificielle et du *machine learning*. Ce sont des

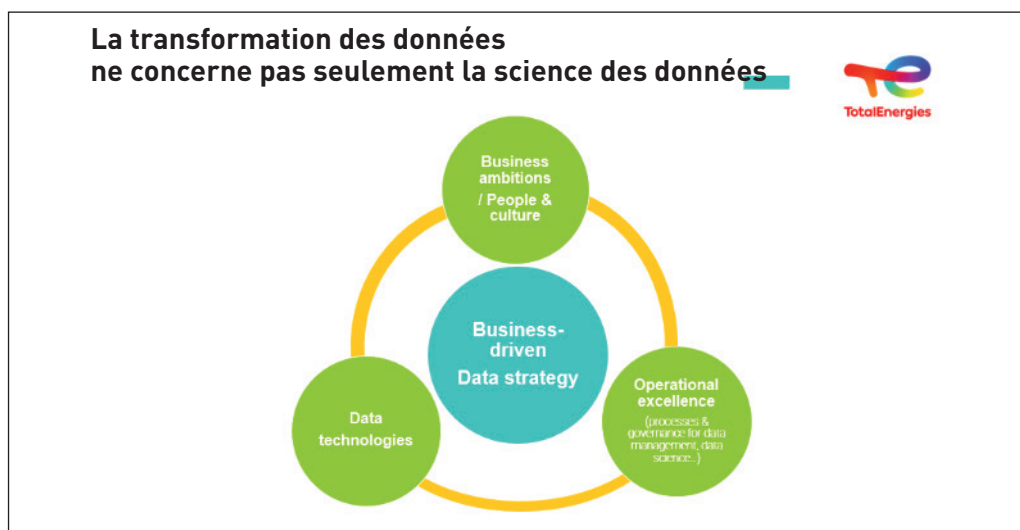


Figure 5

grands chantiers menés avec la direction informatique de la compagnie.

Data management

Ensuite, il y a l'excellence opérationnelle à avoir, rien qu'au niveau de la donnée. On peut faire de la super data science mais si la donnée est de mauvaise qualité, ou n'est pas maîtrisée, on prendra nécessairement des mauvaises décisions quelle que soit la qualité des scientifiques qui font des modèles. Donc ce sont des sujets qu'il faut travailler avec les métiers pour mettre en place des processus, des outils, des bonnes pratiques pour garder la maîtrise de la qualité de nos données. C'est vrai pour la data science, mais c'est utile de façon générale pour la prise de décision au quotidien : pas de bonnes données, pas de bonnes décisions. C'est un vrai enjeu de toujours progresser sur ce point.

Vision métier et transformation

Le troisième axe, je l'ai appelé vision métier et transformation. L'objectif est de donner à tous les membres de la compagnie la volonté d'aller toujours plus loin dans l'utilisation de la donnée.

Il faut donner l'envie et la capacité à tout le monde, d'être plus autonome pour utiliser la donnée pour mieux décider.

Il faut aussi transformer les façons de faire avec les données. Dans une société comme TotalEnergies, beaucoup de gens n'ont pas attendu les *data scientists* et les informaticiens pour manipuler les données, pour modéliser des problèmes. Cependant les gens qui viennent du *machine learning* apportent un regard un peu différent sur la modélisation d'un problème.

Il faut donc faire se rencontrer les cultures, il faut savoir comment faire cohabiter des gens

qui vont partir avec juste de la donnée et un peu moins de connaissances avec ceux qui modélisent plutôt un problème avec simulation numérique ou avec une connaissance vraiment phénoménologique. Il faut arriver à trouver les ponts pour qu'ils travaillent ensemble et, finalement, qu'ils regardent différemment les problèmes pour mieux décider ensemble. C'est un des enjeux importants et passionnants du développement de l'IA.