La data science comme méthode expérimentale augmentée

Table des matières

La data science comme méthode expérimentale augmentée..... 1
De De Groot à une modélisation interprétable des systèmes complexes..... 3
1. La data science comme prolongement moderne de la méthode expérimentale..... 3
2. La méthode de De Groot comme ossature de la démarche..... 4
3. L'historique : replacer les données dans une histoire scientifique..... 5
4. L'observation : faire parler les données..... 5
5. L'induction : le moment créatif de la méthode..... 6
6. Le rasoir d'Occam et la justification des modèles réduits..... 7
7. La déduction : tirer les conséquences de l'hypothèse..... 7
8. Les tests : revenir aux données..... 8
9. L'évaluation : stabiliser le modèle et construire un simulateur..... 9
10. L'application : donner une finalité au modèle..... 9
11. Tableau général de synthèse pour le site..... 10
12. Commentaire du tableau..... 11
13. Ce que cette approche apporte à la méthode..... 12
14. Reformulation enrichie de la méthode..... 12
La data science comme méthode expérimentale..... 13
Conclusion..... 13

De De Groot à une modélisation interprétable des systèmes complexes

L'intérêt de mobiliser la méthode empirique d'**A. D. de Groot** est qu'elle donne une profondeur épistémologique à votre démarche. Votre approche ne se réduit pas à une pratique personnelle de la data science ni à une intuition artisanale construite au fil de votre expérience. Elle s'inscrit dans une tradition plus ancienne et plus fondamentale : celle de la **construction scientifique d'un modèle à partir de l'observation**.

Ce point est essentiel pour votre site. Il permet de montrer que votre méthode ne consiste pas seulement à coupler data science et physique de la complexité. Elle consiste à utiliser la data science comme un **instrument moderne de la méthode expérimentale**, et la physique de la complexité comme un **langage de modélisation interprétable**.

Autrement dit, vous ne proposez pas une data science de plus. Vous proposez une manière de redonner à la data science sa fonction expérimentale : observer, formuler des hypothèses, déduire des conséquences, tester, évaluer, puis appliquer.

1. La data science comme prolongement moderne de la méthode expérimentale

La méthode expérimentale classique reposait sur l'observation du réel à travers des instruments de mesure, des expériences, des tableaux de résultats, des courbes expérimentales et des raisonnements inductifs.

Aujourd'hui, dans de nombreux domaines, l'observation ne passe plus seulement par l'expérience de laboratoire. Elle passe par des bases de données, des signaux, des capteurs, des images, des séries temporelles, des simulations, des journaux d'événements, des traces numériques et des systèmes d'information.

La data science devient alors le prolongement contemporain de la méthode expérimentale. Elle n'est pas seulement une boîte à outils algorithmique destinée à produire des prédictions. Elle peut devenir le moyen moderne d'observer des systèmes complexes, d'en extraire des régularités, de formuler des hypothèses, puis de construire des modèles interprétables.

C'est une idée forte pour votre site :

> La data science n'est pas seulement une technique de prédiction. Elle peut redevenir une méthode expérimentale appliquée aux systèmes complexes.

Dans cette perspective, les données ne sont pas de simples nombres. Elles sont les **traces visibles d'un système réel**. Un signal radar est la trace d'une interaction entre la mer, l'onde électromagnétique, la géométrie d'observation et le traitement du signal. Une série financière est la trace d'un système économique, psychologique, institutionnel et spéculatif. Des indicateurs projet sont les traces d'une organisation sous contrainte.

L'enjeu n'est donc pas seulement de traiter les données. Il est de comprendre **quel système les produit**.

2. La méthode de De Groot comme ossature de la démarche

La méthode empirique de De Groot peut être résumée par la séquence suivante :

Historique → Observation → Induction → Déduction → Tests → Évaluation → Application

Votre propre méthode était formulée ainsi :

Observer. Relier. Réduire. Modéliser. Interpréter. Transmettre.

Les deux approches ne se contredisent pas. Elles se complètent. De Groot donne la structure scientifique générale. Votre méthode donne la traduction contemporaine dans le champ de la data science, de la physique de la complexité et des modèles réduits.

Méthode de De Groot	Votre méthode	Sens dans votre approche
Historique	**Relier**	Comprendre ce qui a déjà été fait, ouvrir la bibliographie, chercher des analogies transdisciplinaires.
Observation	**Observer**	Examiner les données, les mesures, les signaux, les distributions, les anomalies.
Induction	**Réduire**	Faire émerger une hypothèse structurante : variable d'ordre, variable latente, phénomène dominant.
Déduction	**Modéliser**	Déduire les conséquences logiques de l'hypothèse et construire le modèle réduit.
Tests	**Interpréter / valider**	Vérifier si les conséquences du modèle apparaissent dans les données.
Évaluation	**Modéliser / formaliser**	Stabiliser le modèle, identifier ses paramètres, construire éventuellement un simulateur.
Application	**Transmettre / décider**	Utiliser le modèle pour prédire, comprendre, décider, concevoir ou former.

Ce tableau est fondamental. Il montre que votre méthode artisanale n'est pas une démarche subjective ou isolée. Elle reprend les grandes étapes de la construction scientifique, mais les reformule dans le langage des systèmes complexes.

L'étape **historique** devient la recherche bibliographique et transdisciplinaire.
L'étape **observation** devient l'EDA au sens profond : non pas quelques graphiques, mais une écoute du système.
L'étape **induction** devient la recherche d'une variable d'ordre ou d'un mécanisme dominant.
L'étape **déduction** devient la construction d'un modèle réduit.
Les **tests** permettent de confronter les conséquences du modèle aux données.
L'**évaluation** stabilise le modèle et peut conduire à un simulateur.
L'**application** donne à la démarche sa portée opérationnelle.

Ainsi, votre méthode peut être comprise comme une **méthode expérimentale augmentée par la data science**.

3. L'historique : replacer les données dans une histoire scientifique

Dans une data science trop rapide, on commence souvent directement par les données : on charge un fichier, on nettoie, on visualise, puis on entraîne un modèle.

Votre démarche est différente. Avant même de traiter les données, il faut comprendre l'histoire du problème.

Quels travaux existent déjà ?
Quels modèles ont été proposés ?
Quels résultats sont robustes ?
Quelles limites ont été rencontrées ?
Quels domaines apparentés peuvent éclairer la situation ?
Quelles analogies transdisciplinaires peuvent ouvrir une hypothèse nouvelle ?

Cette étape correspond à ce que De Groot appelle l'**historique**. Dans votre méthode, elle correspond au verbe **relier**.

Relier, ce n'est pas seulement faire une bibliographie. C'est replacer le problème dans une constellation de savoirs. C'est chercher dans d'autres disciplines des systèmes qui présentent des formes comparables : turbulence, finance, biologie, réseaux, organisations, transitions de phase, diffusion/réaction, fractalité, mémoire longue.

Dans l'exemple du fouillis de mer radar, cette étape est très claire. Avant de proposer un modèle, il faut connaître les travaux sur la prévision de houle, les spectres spatio-temporels de la mer, les modèles de rétrodiffusion, les mesures radar dans différentes bandes de fréquence, les conditions météo, les modèles de clutter et les analyses non linéaires de signaux radar.

La donnée n'arrive donc jamais seule. Elle arrive avec une histoire de mesures, de modèles, de controverses, d'approximations, d'échecs et de résultats empiriques.

C'est pourquoi une phrase importante pour votre site pourrait être :

> Observer, ce n'est pas seulement regarder les données. C'est aussi replacer ces données dans l'histoire scientifique du problème.

4. L'observation : faire parler les données

L'observation est la première véritable confrontation avec le réel. Dans votre méthode, elle correspond à l'EDA, mais une EDA enrichie.

Il ne s'agit pas seulement de produire des histogrammes, des courbes ou des matrices de corrélation. Il s'agit d'écouter le système.

Observer, c'est chercher :

les régimes,
les ruptures,
les cycles,
les corrélations longues,
les distributions,
les queues lourdes,
les événements extrêmes,
les signaux faibles,
les échelles pertinentes,
les interfaces actives,
les variables cachées.

Dans l'exemple du clutter de mer, l'observation porte sur des campagnes d'enregistrement de données radar à faible incidence. L'analyse statistique de l'énergie rétrodiffusée montre des lois de répartition à queue lourde : les grandes valeurs ne sont pas aussi rares que dans un cadre gaussien classique. Les histogrammes sont comparés à plusieurs lois usuelles : lognormale, K, Rayleigh. Mais aucun de ces modèles ne paraît pleinement satisfaisant.

C'est précisément là que l'observation devient intéressante. Elle ne confirme pas simplement un modèle déjà connu. Elle fait apparaître un écart, une anomalie, une difficulté. Or, dans une démarche scientifique, une anomalie n'est pas seulement un problème. Elle peut être le signe qu'un mécanisme plus profond est à l'œuvre.

Dans votre méthode, les grandes valeurs du clutter ne doivent donc pas être traitées comme de simples aberrations. Elles deviennent des signatures possibles de la structure physique du système : facettes de mer, fortes pentes, masquage, glints, texture intermittente, géométrie ultra-rasante.

5. L'induction : le moment créatif de la méthode

L'induction est probablement l'étape la plus importante pour comprendre votre originalité.

À partir de ce qui est observé, on cherche l'hypothèse la plus plausible. On ne prouve pas encore. On formule une explication candidate.

Dans le cas du fouillis de mer, le raisonnement inductif pourrait être formulé ainsi :

Les données radar présentent des queues lourdes.
Les grandes rétrodiffusions ne sont pas rares.
Les modèles classiques ne suffisent pas à expliquer complètement cette structure.
Le radar observe la mer à faible incidence.
Dans certaines configurations, des vagues ou des facettes de mer peuvent se comporter comme des miroirs.
Les fortes rétrodiffusions pourraient donc venir de facettes fortement inclinées.
Il devient alors plausible que la distribution des pentes de mer joue un rôle central.
Si cette distribution est à queue lourde, elle peut expliquer la fréquence des grands retours radar.

C’est ici que l’on passe de l’observation brute à une hypothèse structurante.

Dans votre approche, cette étape correspond à la recherche d’une **variable latente**, d’une **variable d’ordre** ou d’un **mécanisme dominant**.

Dans le cas radar, la variable latente peut être la pente locale, ou la texture qui traduit l’effet combiné des pentes, du masquage et de la rugosité.

Dans un marché financier, la variable latente peut être la confiance, la volatilité lente ou un régime de marché.

Dans un projet industriel, elle peut être une tension organisationnelle ou un rapport entre sollicitation et capacité de réponse.

Cette étape inductive est aussi celle où s’exprime le mieux votre posture d’artisan. Vous ne vous contentez pas d’ajuster une loi. Vous cherchez le mécanisme plausible. Vous essayez de comprendre ce que les données disent du système qui les produit.

On pourrait formuler cela ainsi :

> L’induction est le moment où les données cessent d’être seulement observées et commencent à devenir une hypothèse sur le fonctionnement du système.

6. Le rasoir d’Occam et la justification des modèles réduits

Votre approche privilégie les modèles réduits, à peu de paramètres significatifs. Ce choix n’est pas seulement pratique. Il est profondément lié au rasoir d’Occam.

Le rasoir d’Occam consiste à privilégier l’hypothèse la plus simple capable d’expliquer les observations. Cela ne signifie pas choisir une explication simpliste. Cela signifie éviter d’ajouter des paramètres ou des mécanismes inutiles.

Dans les systèmes complexes, cette idée est décisive. On ne peut pas tout modéliser. On ne peut pas représenter tous les degrés de liberté. Il faut donc identifier les mécanismes dominants.

Un modèle réduit n’est pas un modèle pauvre. C’est un modèle qui conserve l’essentiel.

Dans le cas du radar, au lieu de construire un modèle empirique très paramétré, vous cherchez une chaîne physique lisible :

surface fractale
↓
distribution des pentes
↓
facettes brillantes
↓
masquage
↓
texture
↓
signal radar

Cette chaîne est forte parce qu'elle relie chaque étape à un mécanisme physique. Elle ne cherche pas seulement à reproduire une statistique. Elle cherche à expliquer pourquoi cette statistique apparaît.

C'est exactement ce qui distingue un modèle interprétable d'un modèle opaque.

Une formulation possible pour votre site serait :

> Le modèle réduit est l'expression moderne du rasoir d'Occam appliqué aux systèmes complexes : conserver peu de paramètres, mais leur donner un sens physique fort.

7. La déduction : tirer les conséquences de l'hypothèse

Une hypothèse devient scientifique lorsqu'elle produit des conséquences observables.

C'est l'étape de la déduction.

Une fois l'hypothèse posée — par exemple, les fortes rétrodiffusions viennent de facettes de mer fortement inclinées dont la distribution des pentes est à queue lourde — il faut demander :

Si cette hypothèse est vraie, que doit-on observer ?

Dans le cas radar, plusieurs conséquences apparaissent naturellement.

Si les fortes rétrodiffusions viennent de facettes fortement inclinées, alors le signal doit être intermittent.

Si les pentes suivent une loi à queue lourde, alors l'amplitude radar doit aussi présenter une queue lourde.

Si l'état de mer augmente la rugosité, alors les événements extrêmes doivent devenir plus probables.

Si l'incidence est rasante, alors le masquage doit devenir un mécanisme structurant.

Si la surface est multi-échelle, alors le spectre Doppler doit porter la trace de cette structure.

Cette étape est essentielle, car elle distingue une simple analogie d'un modèle scientifique. Une analogie dit : "cela ressemble à". Une hypothèse déductive dit : "si cela est vrai, alors on doit observer ceci".

La déduction donne donc au modèle sa puissance explicative.

8. Les tests : revenir aux données

La beauté d'une hypothèse ne suffit pas. Il faut revenir aux données.

C'est l'étape des tests. Elle permet d'éviter le piège d'une théorie élégante mais déconnectée du réel.

Les tests peuvent prendre plusieurs formes :

comparaison aux distributions observées,
comparaison aux modèles empiriques existants,
validation sur des données expérimentales,
simulation de cas typiques,
comparaison de différents états de mer,
vérification du comportement Doppler,
contrôle des queues de distribution,
analyse des événements extrêmes.

Dans l'exemple radar, les tests peuvent s'appuyer sur les données disponibles, mais aussi sur les lois empiriques de rétrodiffusion de la littérature. On peut comparer les distributions d'amplitude, les coefficients de rétrodiffusion, l'élargissement Doppler, ou encore l'évolution du clutter selon l'état de mer.

Ce retour aux données est essentiel pour crédibiliser votre méthode auprès d'un public d'ingénieurs ou de data scientists. Il montre que votre démarche ne consiste pas seulement à proposer une lecture physique séduisante. Elle consiste à produire des hypothèses testables.

---

9. L'évaluation : stabiliser le modèle et construire un simulateur

Dans la méthode de De Groot, l'évaluation correspond au moment où l'on stabilise le modèle, où l'on construit les éléments théoriques, où l'on vérifie sa cohérence et son intérêt.

Dans votre démarche, cette étape correspond souvent à la construction d'un **simulateur**.

Le simulateur est un objet essentiel. Il matérialise le modèle. Il permet de transformer l'hypothèse en expérience numérique.

Dans le cas du fouillis de mer, le simulateur peut intégrer :

la surface de mer interceptée par le radar,
les points brillants associés à des facettes,
la loi de répartition des pentes,
le paramètre de forme,
le paramètre d'échelle,
la vitesse du vent,
la configuration géométrique du radar,
l'incidence,
la forme de l'onde incidente,
le mécanisme de masquage,
la texture du clutter,
le Doppler.

Ce simulateur ne se contente pas de générer un bruit artificiel. Il produit un signal à partir d'hypothèses physiques interprétables. Il devient donc une forme d'expérience numérique.

Une phrase forte pour votre site serait :

> Un simulateur physiquement informé est une expérience numérique construite à partir d'hypothèses interprétables.

Cette idée relie très bien data science, physique de la complexité et ingénierie.

10. L'application : donner une finalité au modèle

La dernière étape est l'application.

Un modèle n'a pas seulement pour fonction d'être élégant ou explicatif. Il doit aussi permettre d'agir, de prédire, de décider, de concevoir ou de transmettre.

Dans l'exemple radar, l'application peut concerner :

le tracking de cibles de faible pouvoir réflecteur,
la détection en environnement naval,
la mise au point d'algorithmes,
l'analyse des performances,
la prédiction de configurations difficiles,
la distinction entre cible et clutter,
l'amélioration de la robustesse des traitements.

Le modèle permet alors de mieux comprendre le fond marin radar, de simuler des situations réalistes, d'anticiper les cas défavorables et d'améliorer la performance des algorithmes.

Dans d'autres domaines, la même logique peut s'appliquer.

En finance, le modèle peut aider à comprendre les crises, les changements de régime ou les queues lourdes.

En projet industriel, il peut aider à anticiper les dérives, les seuils de rupture ou la propagation des risques.

Dans les organisations, il peut permettre de lire les tensions, les blocages, les résistances ou les capacités de transformation.

Dans les territoires, il peut aider à penser la résilience face aux contraintes économiques, climatiques ou sociales.

11. Tableau général de synthèse pour le site

Le tableau suivant peut devenir une pièce centrale de votre page. Il relie la méthode empirique, la data science, la physique de la complexité et l'exemple radar.

Étape	Question scientifique	Traduction data science	Traduction physique de la complexité	Exemple radar
Historique	Que sait-on déjà ?	Recherche bibliographique, état de l'art, modèles existants.	Identification des familles de phénomènes connues.	Houle, spectres de mer, rétrodiffusion, chaos du clutter.
Observation	Que montrent les données ?	Visualisation, statistiques descriptives, nettoyage, distributions.	Mise en évidence d'un phénomène dominant.	Données radar à faible incidence, queues lourdes de l'énergie rétrodiffusée.
Induction	Quelle hypothèse explique le phénomène ?	Choix d'hypothèses, modèles candidats.	Variable latente, variable d'ordre, mécanisme dominant.	Pentes de vagues à queue lourde, facettes brillantes, loi de Lévy.
Déduction	Quelles conséquences doivent apparaître ?	Prédictions qualitatives, expériences complémentaires.	Intermittence, masquage, texture, Doppler, amplitude heavy-tail.	Les facettes fortes doivent produire glints, queues lourdes et élargissement Doppler.
Tests	Les conséquences sont-elles observées ?	Validation sur données, comparaison à modèles.	Confrontation aux lois empiriques et aux simulations.	Comparaison aux distributions, coefficients de rétrodiffusion, données expérimentales.
Évaluation	Quel modèle stabiliser ?	Construction du modèle, simulateur, paramètres.	Modèle réduit interprétable à paramètres physiques.	Simulateur de clutter avec facettes, pentes, incidence, vent.
Application	À quoi sert le modèle ?	Prédiction, classification, décision, performance.	Compréhension des régimes et des seuils.	Détection/tracking de cibles faibles en environnement naval.

Ce tableau montre que la data science n'est pas seulement un ensemble de techniques.
Elle intervient à toutes les étapes de la méthode expérimentale.

Elle aide à observer.
Elle aide à formuler des hypothèses.
Elle aide à tester.
Elle aide à simuler.
Elle aide à appliquer.

Mais elle ne remplace pas le raisonnement. Elle le prolonge.

12. Commentaire du tableau

L'étape **historique** rappelle que toute modélisation commence par un héritage. On ne construit pas un modèle à partir de rien. On s'appuie sur des travaux antérieurs, mais on cherche aussi leurs limites. C'est là que votre ouverture transdisciplinaire devient importante : une difficulté rencontrée dans un domaine peut parfois être éclairée par une structure connue dans un autre domaine.

L'étape **observation** est celle où les données sont interrogées. Mais l'objectif n'est pas seulement descriptif. Il s'agit de faire émerger un phénomène dominant. Dans le cas radar, ce phénomène est l'apparition de grandes valeurs de rétrodiffusion plus fréquentes que prévu. Dans d'autres domaines, cela pourrait être une volatilité persistante, une rupture de régime, une propagation de tension ou un comportement critique.

L'étape **induction** est le cœur créatif. Elle consiste à proposer une hypothèse explicative. Dans le cas radar, l'hypothèse est que les grandes valeurs viennent de facettes de mer fortement inclinées, dont la distribution des pentes est à queue lourde. Cette hypothèse transforme une observation statistique en mécanisme physique plausible.

L'étape **déduction** oblige à tirer les conséquences de l'hypothèse. Si les pentes sont à queue lourde, alors la texture radar, l'amplitude et le Doppler doivent porter la trace de cette structure. C'est ce passage qui donne au modèle sa cohérence.

L'étape **tests** confronte ces conséquences au réel. C'est là que l'on compare aux données, aux distributions, aux lois empiriques, aux simulations ou aux observations de la littérature. La validation ne signifie pas que le modèle est définitivement vrai, mais qu'il est cohérent, utile et suffisamment robuste pour être exploité.

L'étape **évaluation** stabilise le modèle. Elle identifie les paramètres significatifs, leurs domaines de validité et leurs limites. Dans votre approche, cette étape conduit souvent à un simulateur, c'est-à-dire à une expérience numérique où les hypothèses deviennent manipulables.

Enfin, l'étape **application** donne au modèle sa finalité. En radar, il s'agit d'améliorer la détection et le tracking. En finance, il pourrait s'agir de mieux comprendre les crises. En projet industriel, de repérer les tensions et les risques. En organisation, de lire les blocages et les capacités de transformation.

13. Ce que cette approche apporte à la méthode

Cette synthèse apporte quatre contributions majeures.

Premièrement, elle donne une **légitimité scientifique** à votre démarche. Vous ne proposez pas une méthode marginale, mais une actualisation de la méthode expérimentale dans l'âge des données.

Deuxièmement, elle clarifie le rôle de la data science. La data science n'est plus réduite à l'algorithmique ou au machine learning. Elle devient l'ensemble des moyens permettant

d’observer, d’induire, de déduire, de tester et de valider des hypothèses sur des systèmes complexes.

Troisièmement, elle justifie votre usage des modèles réduits. Le modèle réduit n’est pas un choix esthétique. Il découle du rasoir d’Occam et de la recherche de paramètres interprétables.

Quatrièmement, elle donne à votre site une structure pédagogique très forte. Chaque application pourra être racontée selon le même cycle : historique, observation, induction, déduction, tests, évaluation, application.

Cela permet d’unifier des domaines très différents : radar, finance, projets industriels, organisations, territoires, résilience.

14. Reformulation enrichie de la méthode

Votre méthode initiale était :

Observer. Relier. Réduire. Modéliser. Interpréter. Transmettre.

La méthode de De Groot permet de l’enrichir en profondeur.

Une version plus explicite serait :

Observer. Relier. Induire. Réduire. Modéliser. Tester. Interpréter. Transmettre.

Mais pour un bandeau de site, cette formulation est peut-être trop longue.

Il est préférable de conserver la formule simple :

Observer. Relier. Réduire. Modéliser. Interpréter. Transmettre.

Puis d’expliquer dans la page “Méthode” que cette chaîne recouvre les étapes de De Groot :

Verbe de votre méthode	Correspondance avec De Groot	Sens
Observer	Observation	Examiner les données et mettre en évidence le phénomène dominant.
Relier	Historique	Replacer le problème dans son état de l’art et ses analogies transdisciplinaires.
Réduire	Induction	Faire émerger une hypothèse structurante et une variable d’ordre.
Modéliser	Déduction / Évaluation	Construire un modèle réduit et en tirer les conséquences.
Interpréter	Tests / Évaluation	Vérifier, comprendre les paramètres, lire les régimes et les seuils.
Transmettre	Application	Utiliser le modèle pour décider, concevoir, former ou expliquer.

Ce second tableau est important, car il relie directement votre bandeau de site à la logique scientifique sous-jacente.

La data science comme méthode expérimentale

Cette démarche prolonge la méthode empirique classique dans le monde contemporain des données. Dans l'esprit d'A. D. de Groot, toute construction scientifique passe par une séquence : historique, observation, induction, déduction, tests, évaluation et application.

La data science peut être comprise comme le vecteur moderne de cette démarche expérimentale. Elle permet d'observer les systèmes à travers leurs données, de visualiser les régimes, de repérer les anomalies, de mettre en évidence les distributions, les corrélations et les échelles. Mais elle ne doit pas s'arrêter à la prédiction.

Dans cette approche, les données sont les traces d'un système complexe. L'objectif est donc de comprendre ce qui les produit. L'analyse exploratoire devient une écoute du système. L'induction permet de formuler une hypothèse sur le phénomène dominant. La réduction cherche une variable d'ordre ou une variable latente. La modélisation construit une équation ou un simulateur à peu de paramètres significatifs. Les tests confrontent les conséquences du modèle aux observations. L'interprétation donne un sens aux paramètres, aux régimes, aux seuils et aux transitions. L'application transforme enfin cette compréhension en décision, en conception ou en transmission.

Dans cette perspective, un modèle n'est pas une boîte noire destinée seulement à prédire. Il est une hypothèse organisée sur le fonctionnement du système. Sa valeur tient autant à son pouvoir d'explication qu'à sa capacité de prédiction.

Conclusion

Ce que la méthode de De Groot permet d'extraire est décisif.

Elle donne à votre approche une structure épistémologique solide. Elle montre que votre démarche n'est pas seulement une pratique artisanale de la data science, mais une **actualisation de la méthode expérimentale** à l'époque des données massives, des signaux complexes et des systèmes non linéaires.

Votre formulation devient alors beaucoup plus forte :

> J'utilise la data science comme instrument moderne de la méthode expérimentale, et la physique de la complexité comme langage pour construire des modèles réduits, interprétables et testables.

Cette phrase pourrait devenir l'un des piliers de votre site.

Elle résume votre singularité : vous partez des données, mais vous ne vous arrêtez pas à l'algorithme. Vous cherchez le phénomène dominant, la variable d'ordre, le modèle réduit, l'interprétation physique et la transmission du sens.