

L'IA sous supervision de l'analyste : un cadre pour l'analyse de données d'études assistée par les grands modèles de langage

Walther Oettgen
Sonam, Paris

Jérémy Roch
Sonam, Paris

Stéphane Boucharenc
HPIA

walther.oettgen@sonam-ai.fr jeremy.roch@sonam.com stephane.boucharenc@hpia.fr

Juillet 2026

Résumé

L'intégration des grands modèles de langage (LLM) dans les études de marché et la recherche en sciences sociales est le plus souvent pensée selon deux figures opposées : l'automatisation complète du travail d'analyse, ou le contrôle humain réduit à une validation terminale. Nous proposons un troisième cadre, l'IA sous supervision de l'analyste, dans lequel toute décision qui structure l'interprétation est prise ou ratifiée par ce dernier, le modèle jouant le rôle d'un instrument qui propose, cite ses sources et se corrige. Ce cadre repose sur un argument de nécessité : les données d'une étude sous-déterminent leur interprétation, et l'information qui permet de trancher, contexte de la décision à éclairer, expérience du terrain, pour partie tacite, réside chez l'analyste ; la supervision est le canal par lequel elle entre dans l'analyse. Nous en déduisons quatre principes de conception assortis de critères vérifiables (contextualisation préalable, proposition révisable, traçabilité intégrale, réversibilité), montrons leur opérationnalisation dans des instruments d'analyse dédiés aux études, quels qu'en soient les matériaux (verbatim de questions ouvertes, entretiens, données d'enquête, fonds documentaires), et les illustrons par une comparaison contrôlée sur un corpus réel de 153 verbatims issus d'un test de concept. Onze décisions de supervision, pour un surcoût de dix minutes, y suffisent à restructurer la lecture du corpus : les catégories créées par l'analyste captent 43 % du matériau, les verbatims non codés sont divisés par deux, les freins minoritaires deviennent visibles dans les tris, tandis que la part d'aléa imputable au modèle est bornée à 6 %. Nous discutons enfin les limites du cadre, notamment le coût cognitif de la supervision et le risque de biais de confirmation de l'analyste lui-même.

Mots-clés : intelligence artificielle, grands modèles de langage, études de marché, analyse qualitative, analyse quantitative, supervision humaine, human-in-the-loop.

Title: AI under analyst supervision: a framework for LLM-assisted analysis of research and insights data

Abstract: The integration of large language models (LLMs) into market research and social science research is most often framed through two opposing figures: full automation of the analysis work, or human control reduced to terminal validation. We propose a third framework, AI under analyst supervision, in which every decision that structures the interpretation is made or ratified by the analyst, the model acting as an instrument that proposes, cites its sources and corrects itself. The framework rests on a necessity argument: the data of a study underdetermine their interpretation, and the information that settles the matter (the context of the decision to be informed, field experience, partly tacit knowledge) resides with the analyst; supervision is the channel through which it enters the analysis. From this we derive four design principles with verifiable criteria (prior contextualisation, revisable proposal, end-to-end traceability, reversibility), show how they are operationalised in analysis instruments dedicated to research and insights, whatever the material (open-ended question verbatims, interviews, survey data, document collections), and illustrate them through a controlled comparison on a real corpus of 153 verbatims from a concept test. Eleven supervision decisions, at an extra cost of ten minutes, are enough to restructure the reading of the corpus: the categories created by the analyst capture 43% of the material, uncoded verbatims are halved, minority barriers become visible in the tabulations, while the share of variation attributable to model randomness is bounded at 6%. Finally, we discuss the limits of the framework, notably the cognitive cost of supervision and the risk of the analyst's own confirmation bias.

Keywords: artificial intelligence, large language models, market research, qualitative analysis, quantitative analysis, human oversight, human-in-the-loop.

1 Introduction : un débat mal posé

Les grands modèles de langage (LLM) savent désormais coder un corpus de verbatims, en dégager des thèmes et rédiger une synthèse dans des délais qui ne se comparent plus à ceux du travail manuel. Les études de convergence entre analystes humains et LLM sur des tâches d'analyse thématique montrent des recouvrements réels (De Paoli, 2024), et une offre commerciale abondante promet aujourd'hui l'étude de bout en bout, de la collecte au rapport, sans intervention humaine notable. Ce constat n'est pas discuté ici : sur certaines tâches délimitées, les modèles produisent des sorties dont la qualité apparente rivalise avec celle d'un professionnel des études.

Face à cette offre, la réponse la plus répandue tient en une formule : il faut « garder l'humain dans la boucle » (*human in the loop*)¹. La formule fait consensus, et c'est précisément son problème. Elle ne coûte rien. Éditeurs de plateformes entièrement automatisées, instituts d'études et chercheurs en méthodologie la reprennent à l'identique, alors qu'ils désignent par là des dispositifs incompatibles entre eux. Dans la plupart des implémentations, la « boucle » se réduit à un point de contrôle terminal : le modèle produit, l'humain relit, éventuellement corrige, puis signe. Ce schéma, que nous appellerons validation terminale, maintient formellement un humain dans le processus mais lui retire ce qui fait la valeur de son métier, à savoir la conduite du raisonnement. Dire que l'humain doit rester dans la boucle n'engage à rien. Les vraies questions sont ailleurs : où, quand et avec quels leviers doit-il intervenir ? Ce sont elles qui ne sont pas tranchées.

La littérature académique sur la collaboration entre chercheur et IA, construite pour l'essentiel autour de l'analyse qualitative, a commencé à documenter cette question (Jiang et al., 2021 ; Gao et al., 2023 ; Gebreegziabher et al., 2023 ; Alqazlan et al., 2025), mais elle y aborde le plus souvent la place de l'humain sous l'angle de la fiabilité des sorties : l'analyste intervient pour corriger les erreurs de la machine, réduire les hallucinations, rendre les résultats dignes de confiance. Cette posture, défensive, présuppose que le raisonnement est porté par le modèle et que l'humain en surveille les dérives. Nous défendons le renversement de cette hiérarchie : dans un dispositif d'étude bien conçu, l'intelligence du dispositif est celle de l'analyste, et le modèle est un instrument qui étend sa portée, comme le microscope étend celle de l'œil sans prétendre décider de ce qu'il faut regarder. Ce que les grands modèles changent en profondeur n'est d'ailleurs pas d'abord la vitesse d'exécution : c'est la possibilité, nouvelle, de dialoguer avec la donnée. Là où le professionnel des études produisait la donnée puis la dépouillait, il peut désormais l'interroger, la relancer, la confronter à ses hypothèses ; le cadre proposé organise ce dialogue et le place sous l'autorité de l'analyste. Ce renversement n'est pas un simple parti pris. Nous soutiendrons qu'il répond à une nécessité : un corpus, qualitatif ou quantitatif, sous-détermine son interprétation, et l'information qui permet de trancher entre les lectures possibles ne se trouve pas dans les données mais dans le contexte de l'étude et dans l'expérience de l'analyste. La supervision est le canal par lequel cette information entre dans l'analyse ; s'en passer, ce n'est pas automatiser l'interprétation, c'est la laisser flotter.

Ce renversement n'a de valeur que s'il se traduit en contraintes de conception vérifiables. C'est l'objet de cet article. Nous proposons un cadre, l'IA sous supervision de l'analyste², articulé autour de quatre principes de conception qui nous paraissent aujourd'hui déterminants : **la contextualisation préalable** (le savoir local explicitable, brief, enjeux, connaissance du terrain, entre dans le dispositif avant l'analyse), **la proposition révisable** (toute sortie intermédiaire du modèle est un objet de travail amendable avant qu'elle ne conditionne la suite), **la traçabilité intégrale** (tout enseignement remonte à sa source, verbatim ou calcul reproductible) et **la réversibilité**

1. L'expression vient de l'ingénierie des systèmes automatisés : on parle depuis les années 1950 de *man in the loop* pour désigner l'opérateur humain maintenu dans la boucle de commande d'un système (pilotage, défense aérienne). Elle a ensuite été importée dans le champ de l'apprentissage automatique, puis dans celui de l'encadrement réglementaire de l'IA.

2. Le terme de supervision est pris ici au sens du pilotage exercé par l'analyste sur le processus d'analyse. Il ne doit pas être confondu avec l'apprentissage supervisé (*supervised learning*), qui désigne en apprentissage automatique une famille de techniques d'entraînement des modèles à partir de données étiquetées. Les deux notions sont sans rapport.

(revenir sur une décision d’analyse a un coût faible, faute de quoi la supervision reste théorique). Chacun s’accompagne d’un critère de vérification simple : placé devant un outil donné, on peut établir s’il le respecte ou non. C’est ce qui sépare le cadre proposé des déclarations d’intention sur la complémentarité entre l’humain et la machine, qu’aucune observation ne peut confirmer ni démentir.

Le cadre couvre les études dans toute la diversité de leurs matériaux : verbatims de questions ouvertes, entretiens, données d’enquête, fonds documentaires. Sur les matériaux qualitatifs, la supervision porte sur les catégories : plans de codes, thématiques, concepts. Sur les données d’enquête, elle porte sur les spécifications : quelles questions poser aux données, quels tris à plat ou croisés, quelles analyses multivariées, quels tests, l’exécution des calculs étant confiée à un moteur déterministe et reproductible plutôt qu’au modèle de langage lui-même.

La section 2 situe le cadre dans la littérature sur la collaboration humain-IA en analyse de données d’études et identifie l’angle mort qu’il vient occuper. La section 3 le définit et en formule les quatre principes. La section 4 décrit leur opérationnalisation dans des instruments d’étude conçus par les auteurs, et la section 5 en illustre les effets sur un corpus réel. La section 6 discute les apports et les limites du cadre, dont le coût cognitif de la supervision et le risque de biais de confirmation de l’analyste lui-même.

2 Ce que dit la littérature

2.1 La collaboration humain-IA en analyse : un champ structuré par la question du contrôle

Les travaux sur la collaboration entre chercheurs et systèmes d’intelligence artificielle en analyse de données textuelles précèdent les grands modèles de langage. Dès avant leur diffusion, [Jiang et al. \(2021\)](#) étudiaient les pratiques de travail des chercheurs qualitatifs pour évaluer si ce domaine se prêtait à la collaboration humain-IA ; leur conclusion, souvent citée depuis, est que les chercheurs acceptent l’assistance de la machine à condition de conserver la pleine maîtrise du processus analytique, et que l’incertitude et le désordre apparent de l’analyse inductive ne sont pas des défauts à corriger mais des propriétés du travail interprétatif. Les systèmes conçus dans cette lignée cherchent des compromis entre puissance de traitement et contrôle du chercheur : PaTAT, un prototype de recherche d’aide au codage qualitatif, fait synthétiser par la machine des règles de codage exprimées dans un formalisme lisible, que l’analyste comprend et amende ([Gebreegziabher et al., 2023](#)) ; CoAICoder puis CollabCoder, deux environnements de codage collaboratif dans lesquels le modèle propose des codes que plusieurs codeurs confrontent et arbitrent, montrent au passage un résultat qui doit retenir l’attention : l’assistance d’un modèle partagé accélère l’accord entre codeurs mais peut réduire la diversité des codes produits ([Gao et al., 2023, 2024](#)). L’efficacité se paie d’un rétrécissement de l’espace interprétatif, ce qui est exactement le type d’effet qu’un dispositif d’étude doit savoir détecter et compenser.

Plus récemment, [Alqazlan et al. \(2025\)](#) ont proposé un cadre de théorie ancrée computationnelle explicitement structuré par la présence humaine dans la boucle, motivé par la fiabilité des résultats lorsque des outils d’apprentissage automatique analysent de grands corpus sociaux. Ce cadre est sans doute le plus proche du nôtre dans la littérature académique, et la comparaison est éclairante : la boucle y est pensée d’abord comme une garantie de validité scientifique (le chercheur valide les sorties des modèles pour que les résultats soient dignes de confiance), quand nous la pensons d’abord comme le canal d’entrée d’une information que le modèle ne possède pas. Les deux perspectives sont compatibles, mais elles ne mettent pas le curseur au même endroit : la première corrige la machine, la seconde alimente l’analyse.

2.2 Ce que les modèles savent faire, et où ils échouent

Le second versant de la littérature évalue directement les capacités des grands modèles de langage en analyse. Les explorations d’analyse thématique inductive conduites avec un LLM concluent à la faisabilité de l’exercice : le modèle produit des thèmes plausibles et une partie d’entre eux recoupe ceux d’analystes humains (De Paoli, 2024). Les comparaisons entre analyses manuelles et computationnelles menées avec des praticiens (Gauthier et al., 2023) dressent un tableau cohérent : gain de temps et d’échelle réels, mais perte de la relation intime au corpus qui fonde la confiance de l’analyste dans ses propres conclusions.

Les limites documentées dessinent en creux le cahier des charges d’un dispositif de supervision. Elles sont de quatre ordres. Les hallucinations, d’abord : le modèle peut produire des citations d’appui qui ne figurent pas dans le corpus, défaut rédhibitoire dans un métier où le verbatim fait preuve. L’aplatissement, ensuite : entraîné à produire la continuation la plus probable, le modèle surpondère le consensuel et lisse les cas atypiques, alors même que la valeur d’une étude se loge souvent dans ses marges ; le résultat de Gao et al. (2023) sur la réduction de la diversité des codes en est une manifestation mesurée. La sensibilité et la non-reproductibilité, encore : deux formulations voisines d’une même consigne, ou deux exécutions d’une même consigne, peuvent produire des sorties différentes, ce qui interdit de traiter une sortie unique comme un résultat stable. L’arithmétique, enfin : les capacités de calcul des modèles de langage sont structurellement peu fiables (Frieder et al., 2023), ce qui fonde la règle, posée en section 3.4, de ne jamais leur confier la production d’un chiffre.

Ces limites n’interdisent pas l’usage des modèles ; elles disqualifient une manière de s’en servir, celle qui traite la sortie du modèle comme une conclusion. Chacune trouve sa parade dans un principe de conception : l’hallucination dans la traçabilité, l’aplatissement dans la proposition révisable (l’analyste peut rouvrir ce que le modèle a lissé), la non-reproductibilité dans la réversibilité et dans la délégation du calcul à des moteurs déterministes.

2.3 L’angle mort

De ce parcours se dégage une ligne de partage. La littérature académique outille surtout le contrôle des sorties de la machine ; le discours professionnel, lorsqu’il décrit des professionnels des études devenus superviseurs de systèmes d’IA, revendique le pilotage sans en donner les critères, ce qui interdit de distinguer une supervision effective d’une supervision de façade. Entre les deux manque un cadre qui parte de l’autre bout : non pas « comment fiabiliser ce que produit le modèle » mais « comment mettre le modèle sur la route tracée par l’analyste, en faisant entrer dans l’analyse l’information que lui seul détient et qui en fixe la direction ». Les sections qui précèdent ont établi que cette information existe et qu’elle est en partie tacite ; les critères opérationnels des quatre principes fournissent le test qui manquait au discours professionnel. C’est cette place, à l’articulation d’une littérature qui outille le contrôle et d’une pratique qui revendique le pilotage, que le présent cadre entend occuper.

3 Le cadre proposé : l’IA sous supervision de l’analyste

3.1 Pourquoi la supervision est nécessaire

Avant de définir le cadre, il faut établir qu’il répond à une nécessité et non à une préférence. L’argument tient en trois points.

Le premier est la sous-détermination de l’interprétation par les données³. Un même corpus de

3. Le terme est emprunté à la philosophie des sciences, où la sous-détermination désigne le fait que les données d’observation ne suffisent pas, à elles seules, à départager plusieurs théories concurrentes également compatibles avec elles (thèse dite de Duhem-Quine). Il est pris ici au même sens, appliqué à l’analyse d’étude : un même corpus admet plusieurs lectures cohérentes, et le choix entre elles requiert une information extérieure au corpus.

verbatim, comme un même jeu de tris, admet plusieurs lectures cohérentes : une grille thématique peut être découpée à des grains différents, un même écart statistique peut être lu comme un signal ou comme un artefact de recueil, une récurrence peut être banale dans un secteur et remarquable dans un autre. Rien, dans le corpus lui-même, ne permet de trancher. Ce qui tranche est extérieur aux données : le brief et les enjeux de la décision à éclairer, l'historique du marché et des vagues d'études précédentes, la connaissance que l'analyste a de ce qui est attendu ou surprenant pour ce client, sur ce terrain, à ce moment. Un modèle de langage dispose d'un savoir générique considérable, mais il n'a par construction aucun accès à ce savoir local et situé. Un enseignement d'étude n'a de valeur que rapporté à une décision à prendre, et seul l'analyste détient le contexte de cette décision. La supervision n'est donc ni un garde-fou moral ni un confort professionnel : elle est le seul canal par lequel l'information manquante entre dans l'analyse.

Le deuxième point précise le premier : ce savoir de l'analyste n'est que partiellement explicitable, et ce d'autant plus que son expérience grandit. Une partie peut être formulée à l'avance et transmise au modèle sous forme de contexte (objectifs, cible, historique, consignes) ; nous y revenons ci-dessous. Une autre partie est tacite au sens de [Polanyi \(1966\)](#) : l'analyste expérimenté reconnaît un regroupement mal découpé ou un signal faible pertinent quand il le rencontre, sans pouvoir en énumérer les critères a priori. Ce savoir-là ne peut entrer dans l'analyse qu'au fil de celle-ci, par des interventions situées sur des objets concrets. C'est ce qui interdit de réduire la supervision à un cadrage initial, aussi soigné soit-il.

Le troisième point relève de la responsabilité. Un institut ou un chargé d'études engage sa signature sur des conclusions qui orienteront des décisions. Répondre de ses conclusions suppose de pouvoir reconstituer comment on y est arrivé et d'avoir eu, à chaque étape décisive, la possibilité d'en décider autrement. Une analyse conduite de bout en bout par la machine, sans intervention de l'analyste, ne s'endosse pas ; une analyse supervisée, si.

De cet argument découle une structure temporelle de la supervision, en trois temps, que le cadre devra couvrir intégralement :

1. **Ex ante, la contextualisation (principe P0 ci-après).** L'analyste transfère au modèle, avant toute analyse, le savoir local explicitable : objectifs de l'étude, contexte du client et du marché, cibles, consignes méthodologiques. Ce cadrage est déjà un acte de supervision, et un acte expert : choisir ce qui est pertinent, le hiérarchiser et le formuler mobilise son expérience, et deux analystes ne cadreront pas une même étude de la même façon. Le contexte explicite réduit la sous-détermination ; il ne l'annule pas.
2. **En cours, l'arbitrage.** Pendant l'analyse, l'analyste oriente, amende et tranche sur des objets concrets : c'est ici que le savoir tacite s'exerce. Les principes P1 à P3 présentés en section 3.3 organisent ce dialogue.
3. **Ex post, l'endossement.** L'analyste vérifie, reconstitue et assume les conclusions. La traçabilité en est la condition matérielle.

Ces trois temps sont complémentaires et aucun ne dispense des autres. En particulier, l'objection symétrique à la validation terminale, qui consisterait à dire qu'un contexte initial suffisamment riche rend la supervision en cours superflue, tombe devant le caractère tacite d'une partie du savoir mobilisé : ce qui ne peut être énuméré à l'avance ne peut être injecté dans un cadrage, et ne peut entrer dans l'analyse que par l'arbitrage en cours de route.

3.2 Définition

Nous appelons *IA sous supervision de l'analyste* un dispositif d'analyse dans lequel toute décision qui structure l'interprétation est prise ou ratifiée par l'humain, le modèle ne produisant que des propositions, jamais des conclusions arrêtées. La définition repose sur une distinction entre deux catégories d'opérations. Les décisions structurantes sont celles qui conditionnent la lecture finale des données : le découpage d'un plan de codes, le périmètre d'un regroupement, la qualification d'un écart comme signal ou comme bruit, le choix d'un croisement ou d'un test, la formulation

d'un enseignement. Les opérations d'exécution sont celles qui, une fois la décision prise, appliquent celle-ci à grande échelle : classer dix mille verbatims selon un plan de codes arrêté, calculer les tris demandés, rapprocher des occurrences. Le partage des rôles est alors sans ambiguïté : l'exécution revient au dispositif, qui y apporte une puissance sans commune mesure avec le travail manuel ; les décisions structurantes reviennent à l'analyste, parce qu'elles sont exactement le lieu où s'exerce l'information que lui seul détient (section 3.1).

Deux précisions délimitent la définition. D'abord, ratifier n'est pas subir : la ratification suppose que l'analyste dispose des éléments pour décider autrement, et qu'en décider autrement soit praticable. Un dispositif qui présente ses sorties comme modifiables mais rend la modification coûteuse ou aveugle ne satisfait pas la définition ; c'est tout l'objet des principes qui suivent. Ensuite, la définition porte sur l'architecture du dispositif, pas sur l'usage qui en est fait : un analyste peut toujours choisir de tout accepter sans examen. Le cadre garantit que la supervision est possible et outillée à chaque étape, non qu'elle est exercée ; la responsabilité de l'exercer relève de la déontologie professionnelle, non de la conception logicielle.

3.3 Quatre principes de conception

Les quatre principes découlent de la nécessité établie en section 3.1. Ils ne se distribuent pas un à un sur les trois temps de la supervision, mais les couvrent ensemble : le cadrage *ex ante* appelle une porte d'entrée pour le savoir explicitable (P0) ; l'arbitrage en cours d'analyse appelle des prises sur les propositions du modèle (P1), un moyen d'en vérifier le fondement (P2) et un droit à l'erreur qui rende l'intervention praticable (P3) ; l'endossement *ex post* repose à son tour sur la traçabilité (P2). Chaque principe est formulé avec un critère opérationnel : face à un outil donné, on peut établir s'il le respecte ou non. La figure 1 situe temps et principes sur la boucle supervisée.

P0. Contextualisation préalable. Le dispositif offre, avant toute analyse, des réceptacles structurés pour le savoir local explicitable de l'analyste : objectifs et enjeux de l'étude, connaissance de la cible et du marché, acquis des vagues précédentes, consignes méthodologiques. Ce contexte n'est pas un commentaire d'accompagnement : il conditionne effectivement les générations du modèle, et demeure inspectable et modifiable comme tout objet de travail. Critère opérationnel : existe-t-il un lieu dédié, distinct de la consigne libre, où déposer le contexte de l'étude, et peut-on établir ce que ce contexte change aux sorties ? Un outil dont le contexte se réduit au prompt, ou se perd d'une analyse à l'autre, ne satisfait pas P0.

P1. Proposition révisable. Toute sortie intermédiaire du modèle est un objet de travail que l'analyste peut examiner et amender avant qu'elle ne conditionne la suite de l'analyse. Sur les matériaux qualitatifs, cela concerne le plan de codes avant la codification du corpus, les thématiques avant la note de synthèse, les concepts extraits avant leur consolidation ; sur les données quantitatives, la spécification du plan de traitement (tris à plat, tris croisés, tests de significativité) avant son exécution. Le principe s'oppose directement à la validation terminale : dans un pipeline de bout en bout, les choix intermédiaires existent aussi, mais ils sont pris silencieusement par le modèle et enfouis dans le résultat, où ils ne sont plus ni visibles ni négociables. Critère opérationnel : pour chaque étape structurante, existe-t-il un état intermédiaire affiché, éditable, dont la version amendée par l'analyste est celle qui alimente l'étape suivante ? Un outil qui n'expose que le résultat final, ou qui expose des états intermédiaires en lecture seule, ne satisfait pas P1.

P2. Traçabilité intégrale. Tout enseignement remonte à sa source. Côté qualitatif, chaque thème, chaque concept, chaque citation d'appui est relié aux verbatims dont il procède, consultables en contexte ; côté quantitatif, chaque chiffre est produit par un calcul déterministe et reproductible, exécuté par un moteur de calcul et non par le modèle de langage lui-même, dont l'arithmétique n'offre aucune garantie. La traçabilité sert les trois temps de la supervision : elle permet l'arbitrage en cours d'analyse (on ne peut juger un regroupement qu'en voyant ce qu'il regroupe), elle est la condition matérielle de l'endossement *ex post*, et elle constitue la meilleure défense connue contre les hallucinations, puisqu'une affirmation sans source est immédiatement visible comme telle. Critère opérationnel : en partant de n'importe quelle affirmation du livrable, l'analyste peut-il atteindre

les données sources en un nombre d'actions réduit, sans quitter l'outil ? Une affirmation orpheline suffit à mettre le dispositif en défaut.

P3. Réversibilité. Revenir sur une décision d'analyse a un coût faible. Là où la révisabilité (P1) porte sur les propositions du modèle avant qu'elles ne conditionnent la suite, la réversibilité porte sur les décisions déjà arrêtées, une fois leurs effets visibles. L'analyste peut corriger un code, fusionner deux regroupements, requalifier un signal ou amender une spécification sans invalider le travail déjà accompli ni relancer l'ensemble du traitement. Ce principe est celui qui rend les deux autres effectifs : un droit de regard assorti d'un coût de correction prohibitif est un droit théorique, et l'expérience montre que l'analyste cesse d'intervenir dès lors que chaque intervention lui coûte une relance complète. La réversibilité est aussi ce qui autorise l'exploration, constitutive du travail d'étude : essayer un découpage, le comparer à un autre, revenir en arrière. Critère opérationnel : que coûte à l'analyste un changement d'avis à l'étape k ? Si la réponse est « reprendre à l'étape 1 », le dispositif ne satisfait pas P3.

Les quatre principes sont solidaires. P1 à P3 sans P0 (Contextualisation) privent le dialogue de son cap : on arbitre bien, mais sur une analyse qui ignore l'étude. P1 (Révisabilité) sans P2 (Traçabilité) donne un droit de regard aveugle : on peut amender, mais sans voir sur quoi reposent les propositions. P1 (Révisabilité) sans P3 (Réversibilité) donne un droit de regard rhétorique : on peut amender, mais à un coût qui dissuade de le faire. P2 (Traçabilité) sans P1 (Révisabilité) donne la transparence sans le pouvoir : on voit tout, on ne décide de rien. C'est leur conjonction qui transforme la formule consensuelle de « l'humain dans la boucle » en une architecture vérifiable.

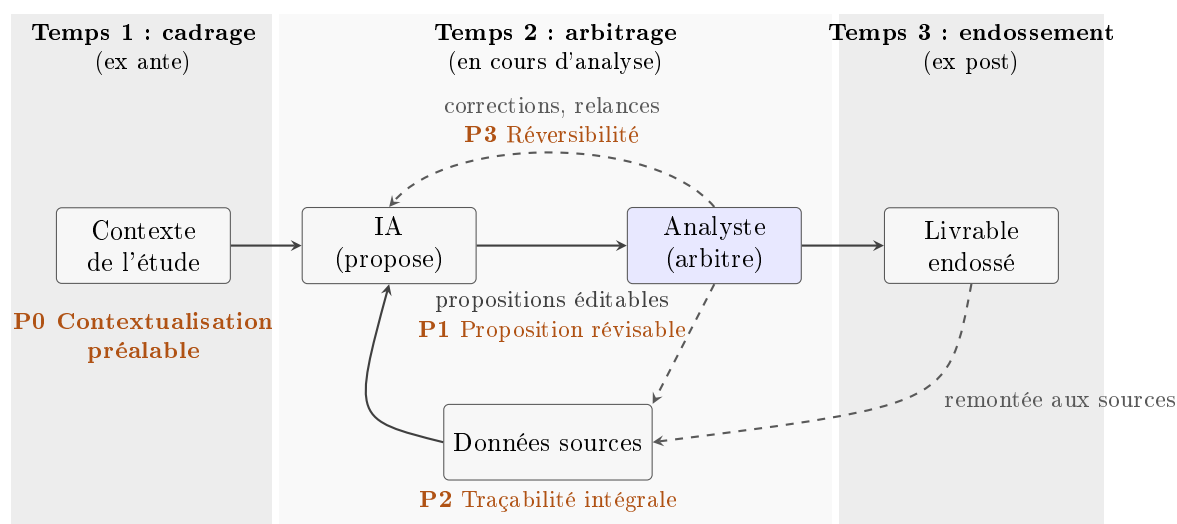


Figure 1 – Les trois temps de la supervision et les quatre principes situés sur la boucle. Le contexte de l'étude entre avant l'analyse (P0) ; en cours d'analyse, les propositions du modèle restent éditables (P1) et revenir sur une décision coûte peu (P3) ; toute affirmation, en cours d'analyse comme dans le livrable final, remonte à ses sources (P2), condition de l'endossement.

3.4 Portée : du qualitatif au quantitatif, et au-delà

Le cadre a été motivé par des exemples qualitatifs, mais rien dans sa définition ne l'y restreint, et son extension au quantitatif est directe. Ce qui change d'un terrain à l'autre est la nature des objets sur lesquels porte la supervision, non sa structure. Sur les matériaux qualitatifs, les décisions structurantes portent sur des catégories : plans de codes, thématiques, concepts, personas. Sur les données quantitatives, elles portent sur des spécifications : quelles questions poser aux données, sur quelles bases et quels sous-échantillons, avec quels tris croisés et quels tests de significativité, à quels seuils. Dans les deux cas, le modèle propose (un plan de codes, une batterie d'analyses pertinentes au regard du brief), l'analyste amende et arrête, puis l'exécution est déléguée : au modèle pour la codification à grande échelle, à un moteur de calcul déterministe pour tout ce qui est numérique.

(a) Pipeline automatique



(b) Boucle supervisée

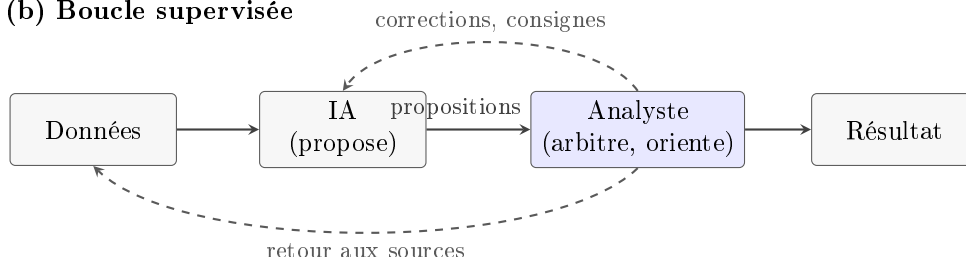


Figure 2 – Deux architectures de la relation entre l’analyste et l’IA. En (a), l’humain intervient en bout de chaîne, sur un produit fini. En (b), l’analyste pilote une boucle : l’IA propose, il arbitre et réoriente, avec un accès permanent aux données sources.

Cette dernière distinction mérite d’être soulignée, car elle est propre au versant quantitatif du cadre. Un modèle de langage peut proposer une analyse pertinente ; il ne doit pas la calculer lui-même, ses erreurs arithmétiques étant documentées et, surtout, non reproductibles. Le partage est donc triple sur le versant quantitatif : l’analyste décide du sens des questions, le modèle aide à les formuler et à en interpréter les résultats, le moteur de calcul les exécute. La supervision porte sur le premier et le deuxième niveau ; la fiabilité du troisième relève de l’ingénierie classique, et c’est précisément pour cela qu’il faut l’y confier. En revanche, une fois les résultats produits (tableaux, graphiques, indicateurs de qualité), les modèles peuvent de nouveau être sollicités pour aider à la rédaction et à la synthèse des conclusions, selon une démarche explicite qui relève de l’expertise de l’analyste.

Cette distinction de matériaux ne doit pas, pour autant, être durcie. Dans un dispositif de dialogue avec la donnée, rien n’interdit de faire dialoguer les matériaux entre eux : éclairer un tri par les verbatims qui lui donnent chair, soumettre une hypothèse née des entretiens à l’épreuve des données d’enquête, confronter un fonds documentaire aux résultats d’un terrain. Le cadre s’applique inchangé à ces allers-retours, parce que la supervision y porte toujours sur le même objet : le sens des questions posées à la donnée, et la responsabilité des réponses qu’on en tire.

4 Opérationnalisation : des instruments d’étude

Un cadre de conception ne vaut que s’il est implémentable. Cette section décrit l’opérationnalisation des quatre principes (P0 à P3), dans un ensemble d’instruments d’étude développés par les auteurs (Sonam AI Études). Le temps ex ante de contextualisation est porté par le principe P0. Les trois instruments sont accessibles en ligne depuis <https://etudes.sonam-ai.fr>. Trois instruments y sont mobilisés : Léon, dédié à la post-codification des questions ouvertes (élaboration du plan de codes, codification du corpus, tris et visualisation des résultats) ; Hippolyte, dédié à l’exploration des matériaux qualitatifs (entretiens individuels, réunions de consommateurs), à l’analyse thématique et à la construction de personas, que l’analyste peut ensuite questionner : l’IA ouvre ici un registre de simulation, le persona construit sur le corpus devenant un interlocuteur que l’on interroge comme on prolongerait un entretien ; Gaston, dédié à l’extraction d’enseignements et de concepts à partir de fonds documentaires d’études. La présentation de la mise en pratique suit les principes plutôt que les outils, afin de montrer que chaque principe admet des implémentations différentes selon la nature de la tâche ; les outils servent ici de cas d’implémentation, et les mécanismes décrits sont transposables à tout dispositif poursuivant les mêmes objectifs.

4.1 La contextualisation en pratique

Le principe P0 (Contextualisation), qui couvre le temps ex ante de la supervision (section 3.1), s'implémente sous la forme de blocs de contexte structurés que l'analyste renseigne avant l'analyse : objectifs et enjeux de l'étude, connaissance préalable du marché et des cibles, segmentations issues de travaux antérieurs, résultats d'études précédentes. Deux choix de conception méritent d'être soulignés. D'une part, ces blocs sont structurés par nature d'information plutôt que réduits à une zone de consigne libre : la structure guide l'analyste dans l'explicitation de son savoir local, précisément parce que cette explicitation ne va pas de soi. D'autre part, chaque bloc est activable ou désactivable pour une analyse donnée : l'analyste décide non seulement du contenu du contexte mais de son usage, et peut comparer une même analyse avec et sans un élément de cadrage, ce qui rend visible l'effet du contexte sur les sorties. Le cadrage cesse ainsi d'être un préambule invisible enfoui dans une consigne pour devenir un objet de travail à part entière, inspectable et versionnable. Une évolution en cours de développement vise à faire analyser chaque élément de contexte par le modèle lui-même, afin d'en produire une synthèse et d'en dégager les apports au regard des matériaux soumis à l'analyse.

4.2 La proposition révisable en pratique

Dans Léon, l'implémentation de P1 (Révisabilité) est frontale : le plan de codes, proposé par le modèle à partir d'un échantillon du corpus et du contexte fourni, est présenté à l'analyste comme un objet éditable (libellés, regroupements, ajouts, suppressions, fusions) avant que la codification ne soit lancée. La codification du corpus complet, opération d'exécution au sens de la section 3, n'applique jamais que le plan arrêté par l'analyste. Dans Hippolyte, les objets intermédiaires sont les thématiques et les personas : les premières sont renommables et recomposables, les seconds sont générés sur la base des axes de segmentation que l'analyste a fournis et activés dans le contexte, plutôt que sur des axes inventés par le modèle. La constitution même du corpus de travail relève du principe : l'analyste peut construire des sous-corpus par filtrage (exclusion d'un locuteur, typiquement l'animateur d'une réunion de consommateurs, seuils de longueur des verbatims), et ces choix de périmètre, qui sont des décisions structurantes s'il en est, lui appartiennent explicitement. Dans Gaston, le périmètre documentaire de chaque analyse transversale est sélectionné par l'analyste, et la question posée au corpus est la sienne, formulée librement.

4.3 La traçabilité en pratique

L'implémentation de P2 (Traçabilité) repose sur une règle de génération commune aux trois instruments : toute production du modèle qui affirme quelque chose du corpus doit citer ses sources, et les consignes de génération interdisent explicitement les citations reconstruites, les verbatims d'appui devant être tirés du corpus. Dans Léon, la traçabilité est structurelle : chaque verbatim reçoit ses codes, et tout chiffre des tableaux de résultats (effectifs, pourcentages, bases) est calculé par le moteur de traitement à partir des codifications unitaires, jamais généré par le modèle ; la distinction entre base répondants et base réponses, classique dans le métier, est portée par le moteur de calcul. Dans Hippolyte, toute production analytique, thématiques, signaux faibles, synthèses, est restituée avec les verbatims qui la fondent, et chaque verbatim peut être replacé dans son contexte d'énonciation : un clic ramène au passage du corpus dont il est extrait, avec ce qui le précède et le suit, de sorte que l'analyste juge sur pièces si une lecture est fidèle ou surinterprétée. Les personas relèvent du même régime de preuve : produits par itération, acceptés, refusés ou relancés par l'analyste, qui en valide le nombre comme la nature, ils peuvent en outre être confrontés aux locuteurs réels du corpus par une mesure de proximité entre leurs verbatims, ce qui donne à une construction synthétique un ancrage vérifiable dans les données. Dans Gaston, les rapports sont générés sous la contrainte de ne s'appuyer que sur les documents actifs du périmètre, et les enseignements renvoient aux documents dont ils procèdent.

4.4 La réversibilité en pratique

L'implémentation de P3 (Réversibilité) tient d'abord à la non-destructivité des opérations : dans Hippolyte, un sous-corpus filtré est un nouvel objet qui coexiste avec le corpus d'origine, de sorte qu'un choix de périmètre s'essaie et se compare au lieu de s'irréversibiliser ; les corpus se renomment, se fusionnent et se rebasculent librement. Dans Léon, l'analyste peut amender le plan de codes et relancer la codification sans reconstruire son projet, et corriger des codifications unitaires. Un mécanisme de Gaston illustre une facette moins évidente du principe : lorsque le périmètre documentaire change après qu'une synthèse a été générée, l'outil signale que l'analyse est devenue obsolète au regard du nouveau périmètre. La réversibilité ne consiste pas seulement à permettre le retour en arrière ; elle consiste aussi à maintenir la cohérence entre les décisions de l'analyste et l'état des analyses, en rendant visible ce qui doit être rejoué. Un dispositif qui laisse coexister silencieusement une synthèse ancienne et un périmètre nouveau fabrique de la fausse traçabilité.

4.5 Bilan

Le tableau 1 récapitule la correspondance entre les temps de la supervision, les principes et leurs implémentations. Deux observations pour conclure cette section. Premièrement, aucune des implémentations décrites n'est techniquement complexe : des états intermédiaires éditables, des sources attachées aux affirmations, des opérations non destructives relèvent de choix d'architecture, pas de prouesses. Si les dispositifs du marché ne les offrent pas toujours, c'est que l'automatisation de bout en bout est plus simple à construire et plus spectaculaire à démontrer, non qu'elle soit plus difficile à dépasser. Deuxièmement, ces choix ont un coût d'usage : ils demandent à l'analyste de s'impliquer là où un pipeline le laisserait spectateur. Ce coût est constitutif du cadre et non un défaut de celui-ci ; la discussion y revient (section 6).

5 Illustration sur un corpus d'étude

Cette section illustre les effets du cadre par une comparaison contrôlée. Le protocole est intégralement spécifié ici ; les valeurs chiffrées seront celles obtenues sur le corpus retenu.

Table 1 – Correspondance entre temps de la supervision, principes et implémentations.

Temps	Principe	Implémentations décrites
Ex ante	P0	Blocs de contexte structurés, activables par analyse
En cours	P1	Plan de codes éditable avant codification ; thématiques et personas recomposables ; sous-corpus par filtrage ; périmètre documentaire choisi
En cours	P2	Citations tirées du corpus ; chiffres calculés par le moteur, jamais générés ; signaux faibles restitués avec leurs verbatims ; ancrage des personas sur les locuteurs réels
En cours	P3	Opérations non destructives ; recodification sans reconstruction ; signalement des analyses obsolètes
Ex post	P2	Remontée aux sources depuis le livrable ; bases et effectifs auditables

5.1 Protocole

Le matériau est un corpus de réponses à une question ouverte issue d'une étude réelle, anonymisé. Il s'agit d'un test de concept, sans dégustation, portant sur un fromage au lait de brebis à texture tendre : le concept était présenté sous forme rédigée, accompagnée d'illustrations du projet de packaging ; la question ouverte recueillait les motifs d'agrément et de désagrément à l'égard

du concept. Le corpus compte 153 verbatims authentiques, conservés tels quels (orthographe, longueurs de 1 à 44 mots, tonalités mêlées), la seule intervention étant l’anonymisation. L’étude a été conduite auprès d’un échantillon représentatif de consommateurs de fromage, non réfractaires au fromage de brebis. Le recueil a été réalisé en salle (hall test), en face à face, le questionnaire étant administré par un enquêteur ; les 153 verbatims correspondent aux 153 répondants. Le protocole a été exécuté avec Léon, l’instrument de post-codification décrit en section 4. Le même corpus est traité dans deux conditions. En condition A, dite automatique, le dispositif enchaîne sans intervention humaine la génération du plan de codes (catégories et sous-catégories), la codification de l’ensemble des verbatims et la production des tris (effectifs, pourcentages sur base répondants, sentiment dominant) ; l’analyste ne voit que le livrable final, conformément au schéma de la validation terminale. En condition B, dite supervisée, le même analyste applique le cadre dans sa composante en cours d’analyse : révision du plan de codes proposé avant codification, examen des thématiques et des signaux faibles avec remontée aux verbatims, corrections et recodifications le cas échéant. L’illustration proposée ne couvre donc pas le temps ex ante de la contextualisation (P0) : l’instrument mobilisé n’offrait pas encore, au moment de cette étude, de bloc de contexte dédié. Cette absence tient à l’état de l’implémentation et non à un caractère facultatif du principe ; l’exigence méthodologique de contextualisation demeure entière ; le savoir local de l’analyste entre ici à travers ses arbitrages en cours d’analyse plutôt que par un cadrage explicite ex ante. La démonstration porte ainsi sur les principes P1 à P3, et l’effet propre du cadrage (P0) reste à établir par un travail ultérieur.

Trois familles d’indicateurs sont relevées. D’abord des indicateurs de structure : composition des plans de codes, décisions de révision consignées au fil de l’eau par l’analyste (regroupements, renommages, créations, recodages unitaires), verbatims non codés. Ensuite des indicateurs de réaffectation : part des verbatims dont l’affectation diffère entre les deux conditions, calculée par un script déterministe d’appariement des deux bases codées, après neutralisation des simples renommages de catégories afin de ne compter que les déplacements réels de sens ; conformément au principe P2, ce script est conservé et rejouable. Enfin des indicateurs de lecture : ce que disent les tris de tête dans chaque condition, et le sort réservé aux verbatims critiques, minoritaires dans un corpus où les avis favorables dominent, comme on peut l’attendre d’un échantillon constitué de consommateurs non réfractaires au fromage de brebis.

5.2 Résultats

Sur la structure d’abord. Les deux plans comptent six catégories, mais douze sous-catégories en A contre quinze en B. L’analyste a consigné onze décisions : quatre portant sur le plan (regroupement des codes relatifs au format et au conditionnement dans une catégorie de réactions à la présentation, renommage de l’« attraction » en « intérêt », jugé plus conforme à ce que mesure un test de concept, création de sous-catégories absentes du plan proposé, dont le goût et la gourmandise) et sept recodages unitaires, portant pour l’essentiel sur des verbatims que la condition A avait laissés sans code. Le résultat le plus net est là : dix verbatims non codés en A, cinq en B, l’analyste ayant su rattacher des formulations elliptiques (« composition », « par le lait de brebis », « fromage brebis ») que le plan initial ne savait pas accueillir.

Sur la propagation ensuite, qui illustre le levier des décisions structurantes. Onze décisions de l’analyste suffisent à réorganiser une large part de la codification, non parce que le dispositif serait instable, mais parce qu’un plan révisé se propage mécaniquement à tout le corpus lors de la codification. La décomposition, calculée par appariement des deux bases après neutralisation des renommages, l’établit : 54 verbatims sur 153 (35 %) conservent une affectation strictement identique ; 60 (39 %) doivent leur écart aux deux catégories créées par l’analyste, l’évaluation de la composition du produit et les attentes concernant le goût, qui captent ensemble 43 % du corpus, un verbatim évoquant le « 100 % lait de brebis » rejoignant la composition parce que cette rubrique existe désormais et non parce que le modèle a changé de lecture ; 13 affectations sont affinées, 13 verbatims non codés sont récupérés, 2 sont perdus. Le résidu d’écarts inexplicables, attribuable

à la non-reproductibilité du modèle, se limite à 9 verbatims, soit 6 % du corpus. La catégorie « comparaison avec d'autres fromages », proposée en A, a été supprimée et ses dix verbatims redistribués vers l'originalité perçue, l'intérêt pour le brebis et les attentes de goût, où ils travaillent au lieu de constituer une rubrique fourre-tout.

Sur la lecture enfin, qui est l'effet le plus décisif. La condition A titre sur une adhésion massive au lait de brebis (59 % des répondants) ; la condition B décompose ce bloc en un intérêt pour le fromage de brebis (35 %) et une évaluation de la composition, le « 100 % lait de brebis » comme argument (33 %), distinction qui change la recommandation puisque l'un renvoie au goût et l'autre à la formulation et à la promesse d'étiquette. Surtout, les verbatims critiques, huit sur 153, sont en B rattachés à des sous-catégories dédiées qui les rendent visibles dans les tris (manque de détails de la présentation, inquiétude sur l'intensité du goût), là où la structure de A les diluait. Le signal le plus concret du corpus concernant le produit lui-même, l'exigence de trous « pour que ça fasse vrai », n'apparaît comme objet de lecture qu'en B, via la sous-catégorie dédiée au concept de fromage à trous.

Sur le coût enfin, entendu ici comme le temps d'analyste mobilisé : quinze minutes en condition A, vingt-cinq en condition B, soit un surcoût de dix minutes pour l'ensemble des onze décisions consignées. Rapporté à ce qu'il change dans le livrable (près de la moitié du corpus rangée sous des catégories nouvelles, les freins rendus visibles, les tris de tête restructurés), ce surcoût est marginal. Le constat ne se généralise pas mécaniquement, un corpus plus vaste ou plus hétérogène appelant davantage d'examen ; mais il suggère que le coût de la supervision croît avec la complexité du plan de codes bien plus qu'avec la taille du corpus, la codification unitaire restant à la charge de la machine.

Table 2 – Comparaison des conditions automatique (A) et supervisée (B) sur le corpus d'illustration (n = 153 verbatims).

	Condition A	Condition B
Catégories / sous-catégories du plan final	6 / 12	6 / 15
Décisions consignées (révision du plan / recodages)	0	4 / 7
Verbatims non codés	10 (6,5 %)	5 (3,3 %)
Affectations identiques (à renommage près)	réf.	54 (35 %)
Écarts liés aux catégories créées par l'analyste	réf.	60 (39 %)
Écarts résiduels (aléa du modèle)	réf.	9 (6 %)
Part du corpus sous les catégories créées en B	0 %	43 %
Temps analyste (min)	15	25

5.3 Lecture

Deux précautions d'interprétation s'imposent, quels que soient les chiffres. Premièrement, la comparaison ne mesure pas la vérité des analyses mais leur écart : elle établit que les décisions de supervision ont des effets matériels sur le livrable, ce qui suffit à réfuter l'idée que la supervision serait un rituel sans conséquence, mais elle ne démontre pas à elle seule que B est meilleur que A, ce jugement restant appréciable sur pièces. C'est précisément pour permettre ce jugement sur pièces que les exemples qualitatifs (affirmations sans appui, signaux requalifiés) accompagnent les comptages. Deuxièmement, un seul corpus ne fonde pas une généralisation ; l'illustration vaut démonstration de faisabilité et invitation à la réplication, non preuve. Troisièmement, la condition B étant une nouvelle exécution du dispositif, une part des écarts de codification unitaire relève de la non-reproductibilité des modèles plutôt que des décisions de l'analyste ; la décomposition présentée la borne toutefois à 6 % du corpus, l'essentiel du mouvement étant la conséquence mécanique du plan révisé. On notera enfin, par honnêteté, que la supervision n'a pas tout rattrapé : au moins un verbatim critique reste, en B, rangé sous un libellé de sous-catégorie à valence positive qui le dessert. La supervision réduit les défauts du livrable ; elle ne les annule pas, et c'est bien pour cela que la traçabilité, qui a permis de repérer ce cas, doit lui survivre dans le livrable final.

6 Discussion

6.1 Apports

Le cadre proposé apporte trois choses. La première est un déplacement de la justification de la supervision : non plus un impératif de prudence ou une clause de conformité, mais une nécessité informationnelle établie par l'argument de la sous-détermination. Ce déplacement change le statut du contexte, qui cesse d'être un paramètre d'appoint pour devenir le premier acte de l'analyse, et il change le statut de l'analyste, dont l'expérience n'est plus ce que le dispositif tolère mais ce dont il a besoin. La deuxième est un jeu de critères opérationnels. Parce que chaque principe est assorti d'un test, le cadre fournit une grille d'évaluation applicable à tout dispositif du marché : face à un outil, on peut demander où sont les états intermédiaires éditables, comment on remonte d'une affirmation à ses sources, et ce que coûte un changement d'avis. Cette grille sert autant les concepteurs que les acheteurs d'outils, dans un marché où la mention d'un « humain dans la boucle » ne dit rien de l'architecture réelle. La troisième est la structure en trois temps, qui permet de localiser les défauts d'un dispositif : tel outil offre un cadrage riche mais aucune prise en cours d'analyse, tel autre l'inverse.

6.2 Limites

Quatre limites doivent être énoncées sans détour. La première est le coût cognitif de la supervision. Le cadre demande à l'analyste de travailler là où un pipeline le laisserait attendre, et ce travail requiert une compétence : réviser un plan de codes suppose de savoir ce qu'est un bon plan de codes, juger une thématique demande le sens du terrain qui permet de la trouver plate ou féconde, questionner un persona à bon escient est un art d'entretien. Le cadre présuppose l'expertise qu'il prétend valoriser, et son bénéfice est moindre entre les mains d'un utilisateur novice. Il faut en tirer une conséquence de formation plutôt qu'une objection au cadre : si les outils d'analyse deviennent des instruments de supervision, la formation des professionnels des études doit porter sur l'art de superviser, qui ne se confond ni avec l'art d'analyser à la main ni avec celui de rédiger des consignes.

La deuxième limite est symétrique de la première : en redonnant la main à l'analyste, le cadre lui redonne aussi ses biais. Une supervision mal exercée peut devenir l'imposition de conclusions préétablies, le contexte fourni ex ante fonctionnant alors comme une prophétie autoréalisatrice. Le cadre contient ses propres contre-feux, qu'il faut expliciter : la traçabilité joue dans les deux sens (un regroupement forcé par l'analyste contre l'évidence des verbatims est aussi visible qu'une hallucination du modèle), la possibilité de conduire une même analyse avec et sans un bloc de contexte rend mesurable l'effet du cadrage, et les mécanismes d'exploration non guidée, comme la détection de signaux faibles, maintiennent dans le dispositif une source de propositions non sollicitées. Mais aucun de ces mécanismes ne remplace la discipline professionnelle, et le cadre garantit la possibilité de la supervision, non sa qualité.

La troisième limite tient au périmètre de validité. Pour des études à faible enjeu, des corpus très homogènes ou des tâches de pur dénombrement, le surcoût de la supervision en cours d'analyse peut excéder son bénéfice, et une validation terminale soigneuse peut suffire. Le cadre ne prétend pas s'imposer partout ; il prétend que là où l'interprétation compte, la supervision est le seul canal d'entrée de l'information qui la fixe.

La quatrième limite est méthodologique : l'illustration de la section 5 porte sur un corpus et un analyste. La validation empirique du cadre appelle des études multi-corpus et multi-analystes, mesurant notamment la variabilité inter-analystes des livrables supervisés, qui est à la fois une conséquence attendue du cadre (deux praticiens n'apportent pas le même savoir local) et une question ouverte pour le métier.

6.3 Perspectives

Deux extensions se dessinent. La première étend la boucle vers l'amont : si la supervision vaut pour l'analyse, elle vaut pour le recueil, et les entretiens conduits par des systèmes conversationnels posent avec une acuité renouvelée la question des trois temps (cadrer le guide d'animation, arbitrer en cours de terrain, endosser les relances). Les auteurs développent un instrument dédié à ces entretiens ; son examen au regard du cadre fera l'objet d'un travail ultérieur. La seconde étend la boucle vers l'aval dans la durée : le savoir local mobilisé au cadrage d'une étude s'enrichit de chaque étude close, et un dispositif de supervision complet devrait outiller cette capitalisation, en faisant des enseignements validés d'une vague le contexte candidat de la suivante. Une troisième direction concerne la stabilité et la reproductibilité des opérations d'exécution confiées aux modèles : une même phase pourrait être réexécutée à plusieurs reprises, voire soumise à plusieurs modèles distincts, avant d'être synthétisée automatiquement, chaque exécution intermédiaire étant conservée conformément au principe de traçabilité (P2).

7 Conclusion

Cet article est parti d'un constat : l'affirmation selon laquelle l'humain doit rester dans la boucle fait l'unanimité et ne coûte rien, ce qui la rend inutilisable telle quelle pour juger un dispositif d'analyse. Nous avons proposé de la remplacer par un cadre qui, lui, engage : une définition (toute décision structurant l'interprétation est prise ou ratifiée par l'humain), un argument de nécessité (les données sous-déterminent l'interprétation, et l'information qui la fixe réside dans le contexte de l'étude et l'expérience de l'analyste, pour partie tacite), une structure en trois temps et quatre principes de conception assortis de critères opérationnels. Nous avons montré que ces principes s'implémentent avec des moyens d'architecture ordinaires, et que leur application produit des livrables matériellement différents de ceux d'un pipeline automatique.

La portée pratique du cadre tient en une inversion de perspective sur le métier. Si l'analyse assistée par les grands modèles de langage est conçue comme un pipeline, la valeur du professionnel des études décroît avec les progrès des modèles, et son avenir est celui d'un relecteur. Si elle est conçue comme une boucle supervisée, sa valeur croît avec les progrès des modèles, car chaque gain de puissance d'exécution augmente le rendement de l'information que lui seul peut apporter. Un instrument plus puissant ne diminue pas celui qui le pilote ; il rend son pilotage plus décisif. Et il change la nature même du travail : produire puis dépouiller la donnée cède la place à un dialogue avec elle, dont l'analyste tient l'initiative et signe les conclusions. C'est en ce sens, précis et vérifiable, qu'il faut entendre que la vraie intelligence du dispositif est celle de l'analyste, d'autant qu'il est seul à engager sa responsabilité et sa réputation professionnelles, à la différence de l'algorithme sous-jacent ou de l'éditeur du logiciel. Les avertissements en petits caractères dont ces outils sont désormais assortis, rappelant que leurs productions peuvent comporter des erreurs, le soulignent en creux : la responsabilité du résultat n'incombe qu'à celui qui l'endosse.

Références

- Alqazlan, L., Fang, Z., Castelle, M., and Procter, R. (2025). A novel, human-in-the-loop computational grounded theory framework for big social data. *Big Data & Society*, 12(2).
- De Paoli, S. (2024). Performing an inductive thematic analysis of semi-structured interviews with a large language model : An exploration and provocation on the limits of the approach. *Social Science Computer Review*, 42(4). <https://doi.org/10.1177/08944393231220483>
- Frieder, S., Pinchetti, L., Griffiths, R.-R., Salvatori, T., Lukasiewicz, T., Petersen, P., and Berner, J. (2023). Mathematical capabilities of ChatGPT. *Advances in Neural Information Processing Systems (NeurIPS), Datasets and Benchmarks Track*. arXiv :2301.13867.
- Gao, J., Choo, K. T. W., Cao, J., Lee, R. K.-W., and Perrault, S. (2023). CoAICoder : Examining the effectiveness of AI-assisted human-to-human collaboration in qualitative analysis. *ACM Transactions on*

Computer-Human Interaction, 31(1) :6 :1–6 :38.

- Gao, J., Guo, Y., Lim, G., Zhang, T., Zhang, Z., Li, T. J.-J., and Perrault, S. T. (2024). CollabCoder : A lower-barrier, rigorous workflow for inductive collaborative qualitative analysis with large language models. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. ACM.
- Gauthier, R. P., Pelletier, C., Carrier, L.-A., Dionne, M., Dubé, È., Meyer, S., and Wallace, J. R. (2023). Agency and amplification : A comparison of manual and computational thematic analyses by public health researchers. Proceedings of the ACM on Human-Computer Interaction, 7(GROUP). <https://doi.org/10.1145/3567552>
- Gebreegziabher, S. A., Zhang, Z., Tang, X., Meng, Y., Glassman, E. L., and Li, T. J.-J. (2023). PaTAT : Human-AI collaborative qualitative coding with explainable interactive rule synthesis. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. ACM.
- Jiang, J. A., Wade, K., Fiesler, C., and Brubaker, J. R. (2021). Supporting serendipity : Opportunities and challenges for human-AI collaboration in qualitative analysis. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW1) :94 :1–94 :23.
- Polanyi, M. (1966). The Tacit Dimension. University of Chicago Press.