

Quels apports du codage des données qualitatives ?

Analyse et propositions à partir du codage de données relatives aux processus de développement des coopérations en management de projet.

Auteur : Corinne Saubesty – Vallier, ATER à l'IUT de Perpignan, laboratoire DMSP

E-Mail : corinne_saubesty@yahoo.fr

Téléphone : 04 68 21 74 65 // 06 20 59 56 37

Résumé

En management stratégique, si les études de cas qualitatives ont appliqué la méthodologie classique de l'étude de cas selon Yin (1989), on observe depuis quelques années une multiplication des travaux de recherche qualitative qui ont choisi le codage des données comme modalité de traitement des données. Aussi, se pose la question de l'intérêt du codage de données qualitatives.

Le codage des données est présenté dans la littérature comme le moyen de faire face à la richesse des données qualitatives et de conduire une analyse systématique des données (Miles & Huberman, 1985, 1991 ; Van de Ven & Poole, 2002).

Des travaux (notamment Huberman et Miles, 1991 ; Van de Ven & Poole, 1990, 2002) ont développé et appliqué des méthodologies de codage des données adaptées aux études de cas en sciences de gestion et notamment en management stratégique.

Aussi, cet article vise à alimenter le débat portant sur la question de l'intérêt du codage des données qualitatives. Doit-on coder les données qualitatives pour mener une recherche qualitative de qualité ? A cette question, nous apportons des éléments de réflexion à partir d'une expérience du codage de données qualitatives dans les processus de développement des coopérations en management de projet.

A partir de la littérature et d'une application du codage des données dans le cadre d'une recherche portant sur le rôle des outils de gestion sur les processus de développement des coopérations, nous présentons des propositions quant aux apports et aux limites du codage des données qualitatives.

Les différentes phases du processus du codage des données et les résultats obtenus sont présentés. En effet, la compréhension du processus de codage est indispensable pour en comprendre les apports et les limites. Nous proposons un modèle de processus de codage en 5 étapes. Ce modèle montre que le codage des données est un véritable processus d'analyse des données interactif au sens de Huberman & Miles (1991). Nous montrons également que ce processus est un processus d'apprentissage des données et de co-construction de la question de recherche. Ensuite, nous abordons les limites du codage liées à la question de la qualité même du processus de codage. Enfin, pour conclure, nous proposons une voie d'analyse alternative mêlant récit et codage, dont la complémentarité est très importante.

Mots clés : méthodologie, méthode qualitative, codage des données, étude de cas longitudinale, processus.

Quels apports du codage des données qualitatives ?

Analyse et propositions à partir du codage de données relatives aux processus de développement des coopérations en management de projet.

INTRODUCTION

En management stratégique, de nombreuses recherches qualitatives utilisent la méthode de l'étude de cas selon Yin (1989). Dans le cadre des études de cas longitudinales, les données recueillies par le chercheur sont ainsi classiquement réduites et présentées sous la forme de récits. Si les travaux de Yin (1984, 1989) ou Pentland (1999) vantent les mérites du texte narratif, à l'inverse, d'autres travaux (notamment Huberman et Miles, 1985, 1991 ; Van de Ven & Poole, 2002) remettent en question, depuis une dizaine d'années, l'intérêt et la validité du récit d'un cas. Selon Huberman et Miles (1991),

" Le seul texte narratif est une forme de présentation peu puissante et extrêmement difficile à manier. Il est contraignant pour les analystes parce que dispersé, étalé sur de nombreuses pages et difficile à consulter ; il est séquentiel plutôt que simultané et rend délicate la visualisation parallèle de deux ou trois variables ; il n'est habituellement que vaguement structuré et peut devenir monotone et indigeste. De meilleurs formats de présentation sont essentiels pour aboutir à une analyse qualitative valable. (...) »

De plus, comme le soulignent Van de Ven & Poole (2002), le volume de données qualitatives collectées dans le cadre d'une étude de cas dépasse les capacités de traitement des chercheurs. L'analyse systématique des données passe par une stratégie de simplification des données.

Ces travaux (notamment Huberman et Miles, 1985, 1991 ; Van de Ven & Poole, 1990, 2002) présentent le codage des données comme une solution pour réduire la complexité des données. Ces travaux ont développé et appliqué des méthodologies de codage des données adaptées aux études de cas en sciences de gestion et notamment en management stratégique.

Nous rappelons que le codage des données consiste à attribuer un code à un segment de texte. Ce processus de transformation des données permet de simplifier les données. Bien plus qu'un simple outil de mise en forme et de présentation des données, c'est un outil qui entame le processus d'analyse des données. En effet, cet assemblage organisé d'informations permet de tirer des conclusions (Huberman & Miles, 1985, 1991).

Les travaux sur le codage proposent un ensemble varié et original de mises en forme de données, dont notamment des matrices, des graphiques, des tableaux. Selon ces travaux, la probabilité d'en tirer des conclusions valables et de les vérifier est bien plus forte qu'avec un texte narratif.

Des travaux se sont intéressés plus particulièrement aux méthodes appropriées au traitement des données dans le cadre des processus de changement (Miles et Huberman, 1984, 1991 ; Van de Ven et Poole, 1990, 2002 ; Abbott, 1990 ; Pettigrew, 1990). Ils apportent des conseils pour la conceptualisation, la mesure et les analyses des séquences d'événements et s'appuient sur la méthode du codage des données.

Suite à ces travaux, de nombreuses recherches qualitatives françaises portant sur la compréhension de processus en management stratégique ont adopté le codage des données comme modalité de traitement (Forgues, 1993 ; Mazars-Chapelon, 2001 ; Belmondo, 2002 ; Dameron, 2000 ; et Saubesty, 2003).

Face à la multiplication des travaux de recherche qualitative qui ont choisi le codage des données comme modalité de traitement des données, se pose la question de l'intérêt du codage de données. Deux courants de pensées s'opposent : en tant que modalité d'analyse permettant une présentation synthétique et systématique des données, le codage est devenu pour certains chercheurs un critère de qualité d'une recherche qualitative. A l'inverse, ses détracteurs lui reprochent une absence de visibilité de la qualité des résultats de la recherche.

Ainsi, cet article vise à alimenter le débat portant sur la question de l'intérêt du codage des données qualitatives. Doit-on coder les données qualitatives pour mener une recherche qualitative de qualité ? A cette question, nous apportons des éléments de réflexion à partir d'une expérience du codage de données qualitatives dans les processus de développement des coopérations en management de projet.

Cet article s'articule en deux parties.

Une première partie aborde l'application du codage des données dans le cadre d'une recherche portant sur le rôle des outils de gestion sur les processus de développement des coopérations. En effet, la compréhension du processus de codage est indispensable pour en comprendre les apports et les limites. Les différentes phases du processus du codage des données et les résultats obtenus à partir de ce codage des données sont exposés.

A partir de cette application du codage des données et des réflexions sous-jacentes, une deuxième partie présente des propositions quant aux apports et aux limites du codage des données qualitatives. Ces propositions sont confrontées aux travaux de la littérature.

Enfin, nous concluons sur les apports et les limites du codage de données qualitatives en apportant notre position quant à l'intérêt du codage des données qualitatives.

1. LE CODAGE DES DONNEES : LE CAS DE PROCESSUS DE DEVELOPPEMENT DES COOPERATIONS EN MANAGEMENT DE PROJETS

Cette partie expose une application de la méthode du codage aux données recueillies dans le cadre d'une recherche sur les processus de coopération en gestion de projet. Après avoir présenté la recherche conduite, nous approfondissons le processus de codage engagé.

1.1. PRESENTATION DE LA RECHERCHE

Problématique et questions de recherche

La recherche menée s'est intéressée à la compréhension du rôle des outils de gestion (tableau de représentation, groupe de travail, aide à la décision, cahier des charges,...) dans le développement de la coopération transverse au sein d'une organisation. A partir de la littérature, nous avons retenu une définition assez générale de la coopération transversale comme le processus par lequel des acteurs d'entités différentes interagissent dans le cadre d'un travail commun.

Notre question de recherche se décompose en quatre sous-questions :

- **Quelles sont les phases du processus de coopération ?**
- **Quels sont les outils et leur rôle dans les phases du processus de coopération ?**
- **Quel est le rôle des outils dans le passage d'une phase de coopération à une autre ?**
- **Quelles sont les propriétés des outils qui jouent un rôle sur l'évolution du processus de coopération ?**

Pour répondre à ces questions, nous avons été intégrés pendant près de quatre années au sein de la Direction de la Recherche et de la Technologie de la SNCF. Nous y avons observé deux cas de coopération transversale :

- une structure de type équipe projet pour la rénovation d'un train, observée pendant deux ans ;
- une structure pérenne de type pôle de compétence recherche, chargée de la capitalisation, de la veille et de la génération de projets de recherche sur le confort, observée pendant plus de trois ans.

La démarche méthodologique adoptée est une approche qualitative de type étude de cas longitudinale. Nous avons adopté une posture mixte entre observation classique, participante et intervenante. Trois types de données ont été recueillies :

- le contenu des réunions via la rédaction d'un journal de bord ;
- les documents (comptes-rendus de réunions, travaux du groupe, documents d'organisation, ...) archivés chronologiquement dans des classeurs ;
- les entretiens rétrospectifs menés avec les membres du groupe (près de 80 entretiens, d'une durée de 1h30 à 3 heures, retranscrits intégralement pour les deux cas).

Le volume des données recueillies est donc très important, représentant plusieurs milliers de pages.

Le traitement des données

Le traitement des données a été réalisé en deux étapes.

Nous avons tout d'abord analysé les données selon la méthodologie classique de l'étude de cas longitudinale (Yin, 1984, 1989). Dans ce cadre, nous

La rédaction de récits chronologiques

Selon Pettigrew (1990), le récit est un moyen de retranscrire simultanément « *les contextes, le contenu et le processus de changement avec leurs interconnexions dans le temps* ». Aussi, pour analyser les données recueillies, les deux cas ont tout d'abord été présentés sous la forme de récits chronologiques, que nous avons complétés de sous récits et d'analyses thématiques.

Les récits constitués adoptent un format commun de présentation et de compte-rendu des données afin de fournir un premier élément à la comparaison inter-cas.

Cette première phase d'analyse des données et de rédaction des premiers résultats a abouti à des analyses générales. La richesse des données était telle que nous avons rencontré d'énormes difficultés à en tirer des résultats de recherche.

A l'issue de cette première phase d'analyse, il est apparu nécessaire d'adopter un mode de traitement des données systématique et synthétique : le codage des données. Dans le cas de cette recherche, le codage est donc postérieur au recueil des données.

Il est important de retenir l'importance à accorder à la rigueur du recueil de données qualitatives du fait des éventuels changements de modalités de traitement ou des changements d'unités d'analyse, etc. Dans notre cas, l'enregistrement systématique des entretiens et leur retranscription sur papier, la rédaction d'un journal de bord narrant chaque événement nous a

permis ce changement de modalité de traitement des données. Le recueil systématique des données recueillies nous a permis de recourir au codage des données à posteriori.

1.2. LE CODAGE DES DONNEES

Dans le cas de données longitudinale, le codage est un processus qui consiste à décomposer la variable étudiée en plusieurs sous-variables (*appelées catégories de code*), puis suivre le déroulement du processus dans le temps à travers les dimensions (*appelées codes*) que prennent chacune des sous-variables qui composent le processus de coopération.

A partir de la littérature (Huberman et Miles, 1991 ; etc.) et des travaux utilisant la méthode du codage, nous avons constitué un processus de codage des données adapté au contexte de l'étude de cas longitudinale. Nous présentons ci-après le processus du codage en 5 phases :

1. La préparation du codage des données
2. La définition d'une grille de codage
3. Le codage des données
4. L'analyse de la base des événements
5. Interprétation et validation des résultats

1.2.1. PHASE 1 : LA PREPARATION DU CODAGE DES DONNEES

Le codage des données est un processus long et fastidieux. Une première phase préparatoire correspond à la **sélection des données pertinentes pour construire la base de données qui fera l'objet du codage.**

Cette phase est importante en raison des choix que le chercheur doit effectuer et qui auront un impact sur la qualité et la pertinence du codage. Ces choix concernent le type de donnée à coder (les entretiens, les documents, etc.) ; le choix de l'unité de codage (le mot, la phrase, le paragraphe de sens) et le choix des données à coder (les données en rapport avec l'objet de la recherche).

▪ Le choix du type de donnée à coder

Les données recueillies étant de nature différente (des observations, des données secondaires et des entretiens), le codage doit-il porter sur un type de donnée ou sur l'ensemble des données ? Le statut d'un type de donnée est-il le même par rapport à un autre ?

Dans les travaux existants sur les processus, le codage est effectué soit sur des bases d'entretiens, soit sur des bases construites à partir des notes d'observations et de données

secondaires. Ces travaux ont accordé le même statut aux différents types de données, sachant qu'en cas de désaccord, ce sont les données d'observation qui sont retenues.

Le choix des données à coder est lié à la question de recherche. Dans le cas de cette recherche, le codage seul des entretiens ne permettait pas de retracer finement les processus de coopération. Il est alors apparu plus pertinent de **coder l'intégralité des données collectées** (les entretiens, les observations et les données secondaires) en raison de la question de recherche qui était processuelle. Le rôle des outils pouvait se laisser appréhender au travers d'une analyse transversale. De plus, les acteurs n'avaient pas la même représentation des outils et ils n'avaient pas forcément conscience de leur impact.

- **La sélection des passages pertinents**

Un travail de surlignage des données a permis de réduire le volume des données à traiter. Il s'agit de sélectionner en regard de la question de la recherche les passages pertinents qui feront l'objet du codage. Il s'agit donc d'éliminer les passages sans rapport avec la question de recherche, appelés « déchets » par Huberman et Miles (1991).

- **Le choix de l'unité de codage**

Le codage des données pertinentes nécessite le choix de l'unité d'analyse. Le degré de précision que le codage doit atteindre dépend des questions de recherche. Aussi, la codification peut être conduite au niveau du mot, de la ligne ou du paragraphe de sens.

Pour observer le processus de coopération, nous nous sommes intéressés aux interactions entre les acteurs. Nous avons considéré comme **unité d'analyse** le « **paragraphe de sens** ». Un paragraphe de sens est une phrase ou un groupe de phrases qui fait référence à une idée principale et dans notre cas à une interaction.

La relecture des données pertinentes nous a permis de dégager les paragraphes de sens.

- **La constitution d'une base de données chronologique des incidents**

Les décisions précédentes nous ont permis de constituer une base des données à coder pour chaque cas de coopération. Cette base de données est appelée « base chronologique des incidents ». Un incident est une matière première, une observation empirique.

Les données recueillies ont été réunies et classées par ordre chronologique au sein d'une base de données sous format papier. Les entretiens étant des entretiens rétrospectifs et non des entretiens « sur le vif », seuls les passages pouvant être précisément intégrés chronologiquement et qui apportaient un élément complémentaire aux autres données ont été ajoutés dans la base de données. Afin d'exploiter toute la richesse des entretiens, nous avons

également constitué une base de donnée rassemblant les entretiens rétrospectifs, qui apportent un éclairage complémentaire.

Parmi les outils proposés par Miles et Huberman (1984, 1991) pour présenter des données, nous avons choisi de présenter la base des incidents sous la forme d'une matrice chronologique parce que le temps est une variable centrale dans l'étude du processus de coopération. L'objectif est de comprendre rapidement et aisément ce qui s'est passé. L'unité d'analyse utilisée pour la construction de la base chronologique des incidents est la séquence d'interactions autour d'une même thématique.

Chaque ligne de la matrice correspond à un incident, qui correspond lui-même à une interaction sur un thème donné. Pour chaque séquence d'interaction, nous précisons le thème abordé, les acteurs, les outils utilisés. L'interaction fait référence à la première dimension de la coopération, la relation. La précision de la thématique permet de faire référence à la deuxième dimension de la coopération : les connaissances.

Tableau n°1 : Base chronologique des incidents

Date	Support	Extrait	Thème	Personnes/ Entités

La base de données constituée, qui réunit chronologiquement les données relatives aux cas étudiés, permet une première mise en forme du processus et rassemble dans un tableau de synthèse les informations essentielles au codage. Elle permet de reconstituer finement les processus de coopération observés.

1.2.2. PHASE 2 : LA DEFINITION D'UNE GRILLE DE CODAGE

Le codage des données est effectué selon une grille de codage qui correspond à la liste des catégories et des codes, qui vont permettre de suivre l'évolution du processus de coopération.

La définition de la grille de codage des données est conduite en trois étapes : Le choix d'une méthode d'élaboration des codes, La définition des catégories de codes et l'*émergence des codes*.

- ***Le choix d'une méthode d'élaboration des codes***

Le chercheur doit choisir la méthode d'élaboration des codes la mieux adaptée à sa recherche. Il existe trois méthodes d'élaboration des codes qui diffèrent par le moment où le chercheur établit sa grille de codage par rapport à la phase de recueil des données.

Le codage est dit :

- « préétabli » lorsqu'une liste initiale de codes est conçue avant le recueil des données à partir du cadre conceptuel, des questions de recherche, etc. (on parle de précodification des données) ;
- « émergent » lorsque les codes émergent des données collectées (reprend la l'approche enracinée de Glaser (1978) ;
- et enfin, « établi par un plan général » lorsqu'un plan général de codage indique les grands domaines dans lesquels les codes devront être inductivement conçus. Cette approche est une approche intermédiaire aux deux précédentes. Des auteurs proposent des catégories autour desquelles le plan de codage s'articule : selon Grenier et Josserand (1999), les acteurs qui interviennent, les activités menées et les éléments du contexte. » ; selon Pettigrew (1990), le contexte, le contenu et le processus.

Nous avons choisi la méthode du *plan de codage guidé par un plan général préétabli* pour deux raisons. D'une part, notre recherche est de type exploratoire. D'autre part, comme notre recherche porte sur le rôle des outils de gestion sur le processus de développement des coopérations transversales, nous pouvons qualifier les catégories de codes.

▪ **La définition des catégories**

Les catégories de codes peuvent émerger de la littérature ou des données. Dans notre recherche, les catégories qui nous ont permis de suivre le processus de coopération transversale ont été déterminées à partir de notre définition de la coopération transversale : « *le processus par lequel des acteurs appartenant à des départements ou métiers différents au sein d'une même organisation interagissent dans le cadre de la réalisation d'un travail commun.*»

Cette définition nous a amené à repérer la coopération au travers des 3 catégories de codes définies précédemment dans le tableau n°1.

Tableau n°2 : Les catégories de codes de la coopération

Concepts	Catégories de codes
▪ Des acteurs	▪ Les acteurs internes et externes
▪ Interagissent	▪ Les séquences d'interaction ▪ La dynamique de groupe
▪ Dans le cadre d'un travail commun	▪ Les activités

Nous nous intéressons au rôle des outils sur le processus de coopération, ce qui nous a conduit également à coder les données sur les catégories « Outils » et « Evénements de la vie des outils ».

Tableau n°3 : Les catégories de codes des outils

Concepts	Catégories de codes
<ul style="list-style-type: none"> ▪ Les outils 	<ul style="list-style-type: none"> ▪ L'outil ▪ Les événements de la vie des outils

▪ **L'émergence des codes**

Le codage doit permettre d'identifier les changements sur les concepts retenus. Un codage émergent des données est effectué pour chacune des catégories énoncées. Selon Belmondo (2002), « un codage émergent consiste à repérer des régularités dans les données de terrain et à en faire émerger des catégories qui sont ensuite utilisées comme code. »

Tableau n°4 : Extrait du tableau des codes de la catégorie « Evénements de la vie des outils » présentée en annexe B

CODE	DEFINITION	COMMENTAIRES
FERM	Fermeture d'un outil	La fermeture signifie une formalisation, une cristallisation de l'outil qui ne pourra plus être interprété et modifié par les acteurs. Par exemple, le tableau de représentation du confort a été modifié jusqu'à ce que les membres du pôle soient d'accord sur la représentation du confort. Une fois validé, ce tableau est devenu une représentation « fermée » du confort qui a pu être utilisée dans d'autres tâches.

1.2.3. PHASE 3 : LE CODAGE DES DONNEES

Le codage des données étant en partie émergent, l'étape de codage des données est également celle de l'émergence des codes.

▪ **La constitution d'une base des événements**

Le codage de la base des incidents est réalisé à partir de la grille de codage précédemment définie. *Le codage consiste à attribuer pour chaque incident, un code pour chacune des catégories concernées.*

Chacun des codes attribués à un incident constitue un événement. Un incident peut concerner plusieurs codes. Par exemple, l'incident suivant « l'animatrice informe les membres du pôle de l'arrivée d'un nouveau membre » correspond aux événements Information et Arrivée Acteur interne. Ainsi, un incident est une matière première, une observation empirique, alors

qu'un événement est un construit théorique puisqu'il n'est pas directement observé (Van de Ven & Poole, 1990).

Le codage de la base des incidents permet ainsi de constituer une base chronologique des événements sous format informatique. La base des événements (tableau n°5) correspond à la matrice des incidents à laquelle a été ajoutée la colonne « Code ».

Tableau n°5 : Base chronologique des événements

Code	Date	Support	Extrait	Thème	Personnes/ Entités

▪ **Stabilisation de la grille de codage**

La base des événements définitive est construite en plusieurs étapes, de la même manière que la grille de codage. En effet, la grille de codage définie doit « résister » au codage de l'ensemble des données de chacun des cas pour être stabilisée. Aussi, le codage de la base des incidents est un processus long. Plusieurs allers-retours ont été effectués entre la phase 2 (définition d'une grille de codage) et la phase 3 (codage des données) jusqu'à stabilisation complète de la grille de codage. La grille n'est stabilisée que lorsque le double codage a permis de la valider.

▪ **Fiabilité et validité du codage**

Une fois la grille stabilisée est construite la base des événements. Afin de contrôler la fiabilité et la validité de la grille de codage et du codage effectué, nous avons effectué un double codage, tel qu'il est préconisé par Huberman et Miles (1991).

Le double codage est tout d'abord effectué par le chercheur lui-même. Ainsi, après avoir effectué un premier codage des données à partir de la grille des codes stabilisée, nous avons réalisé un second codage deux semaines après le premier afin de vérifier la fiabilité du codage effectué. Selon Huberman et Miles (1991), il est souhaitable d'obtenir une fiabilité codage « initial-recodage » proche de 80%.

Huberman et Miles (1991) préconisent ensuite de faire coder séparément une partie des données par un chercheur extérieur à la recherche. La comparaison des codes attribués à chaque incident par les deux chercheurs permet de définir la fiabilité du codage. La littérature ne précise pas sur quelle part des données doit être effectué le codage.

Huberman et Miles proposent la formule suivante pour calculer la fiabilité entre codeurs :

$$\text{Fiabilité} = \frac{\text{nombre d'accords}}{\text{Nombre total d'accords + désaccords}}$$

Ils énoncent qu'il ne faut pas s'attendre initialement à plus de 70% de fiabilité entre codeurs.

« Par la suite, l'accord à la fois intra et inter-codeur doit atteindre les 90%. » (p. 108)

Dans le cadre de notre recherche, le double-codage a été effectué sur 30% des données recueillies. Nous avons obtenu une validité de 83%, ce qui constitue une fiabilité satisfaisante.

1.2.4. PHASE 4 : L'ANALYSE DE LA BASE DES EVENEMENTS

Une fois les données codées, le choix du traitement et de la présentation des données pour l'analyse se pose. Les matrices d'événements constituées permettent de repérer les changements qui surviennent sur chacune des catégories retenues. Cependant, elles n'offrent pas une vision synthétique du processus, la matrice de chaque cas s'étalant sur une trentaine de pages.

Différentes méthodes sont proposées dans la littérature pour l'analyse de données codées. Selon que la question de recherche est une analyse transversale ou processuelle, différentes modalités peuvent être utilisées.

Notre recherche porte sur le processus de coopération. Nous devons alors mettre en évidence des régularités dans le processus et identifier puis nommer les séquences et les phases qui composent ce dernier (Pettigrew, 1992). Pour cela, nous avons suivi la méthode d'analyse des données processuelles proposée par Van de Ven et Poole (1990) et Van de Ven (1992), que nous présentons en trois étapes :

- analyse de l'évolution des trajectoires ;
- repérage des phases ;
- caractérisation des phases.

1.2.4.1. Analyse de l'évolution des trajectoires

Pour décrire le processus de coopération et son instrumentation, nous devons analyser **l'évolution des différentes catégories de codes au cours du temps.**

- **Construction d'un tableau des occurrences**

De la même manière que Van de Ven et Poole (1990), nous avons suivi l'occurrence et la non occurrence des différents codes pour répertorier les changements intervenus au fil du temps.

La méthode consiste à construire un tableau comportant en colonne les codes des catégories

du codage et en ligne les incidents. Ce tableau permet de relever pour chaque incident l'occurrence d'un événement (notée 1) ou sa non occurrence (notée 0).

Pour chacun des cas, nous avons obtenu un tableau de 29 colonnes (correspondant aux 29 codes) et d'un nombre de lignes correspondant au nombre d'incidents relevés (469 lignes pour le cas Pôle Confort et 218 lignes pour le cas Projet de rénovation).

Nous avons obtenu pour chacune des catégories des codes des différentes catégories retenues, une suite chronologique d'occurrences ou de non occurrences des codes.

A partir de ce tableau, il est possible de réaliser des calculs statistiques simples de type fréquence afin de définir le nombre de fois où un code apparaît au cours d'une unité d'analyse temporelle et ainsi l'évolution dans le temps des trajectoires de chacune des catégories.

Il est nécessaire de définir l'unité d'analyse temporelle pertinente. Dans notre recherche, nous avons choisi le mois comme unité temporelle (fréquence moyenne des contacts).

La trajectoire d'une catégorie correspond à l'évolution du nombre d'occurrences de chaque code d'une catégorie au cours de l'unité temporelle retenue. Par exemple, le mois d'octobre 2003 correspond à une fréquence de 3 du code Création Outil.

▪ **La définition de la trajectoire de chaque catégorie**

L'évolution des codes d'une catégorie permet de **définir la trajectoire de la catégorie**. L'analyse de **l'évolution des différents codes au cours du temps** nous a permis de procéder à une analyse descriptive du processus de coopération et des outils.

A partir du tableau récapitulatif de l'occurrence totale mensuelle de chaque code de chaque catégorie, nous avons construit un **graphique chronologique de fréquence** des codes pour chaque catégorie. Il est ainsi possible de suivre chaque trajectoire, i.e. l'évolution de chaque catégorie. En effet, la visualisation graphique des résultats a de bonnes propriétés selon Van de Ven & Poole (2002). Elle permet de **définir les trajectoires de chaque catégorie au cours de la période d'observation**.

L'analyse indépendante de chaque trajectoire d'une catégorie donnée permet de définir des périodes homogènes pour chacune d'entre elles. Par exemple, concernant les séquences d'interaction, un histogramme chronologique de la fréquence des différents types d'interaction nous a permis de définir plusieurs périodes homogènes d'interaction.

▪ **Rédaction d'un récit de l'évolution de chaque trajectoire**

Nous avons rédigé un récit pour retracer l'évolution de chaque trajectoire.

1.2.4.2. Le repérage des phases

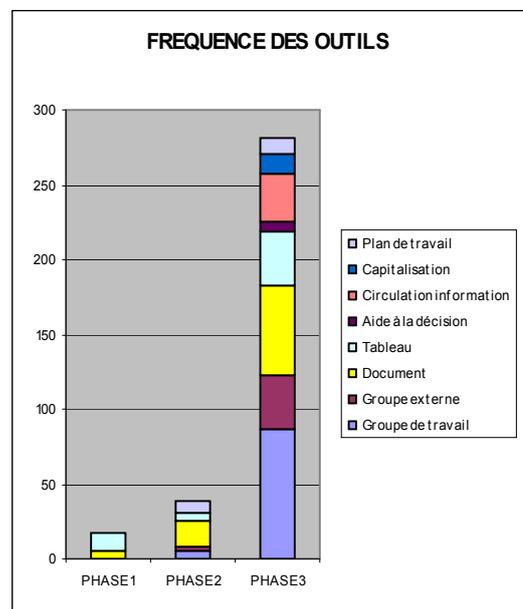
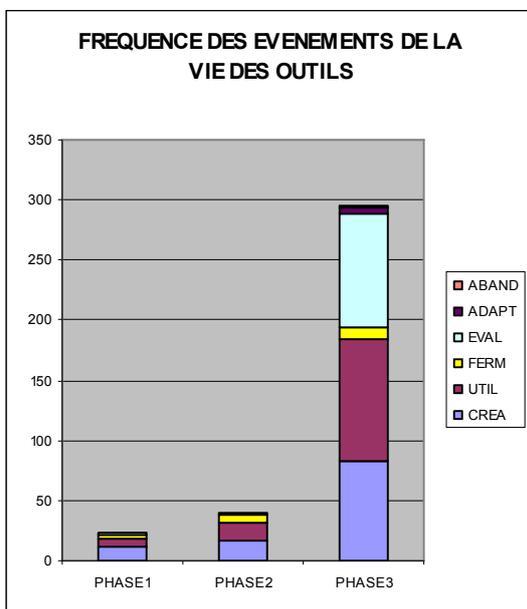
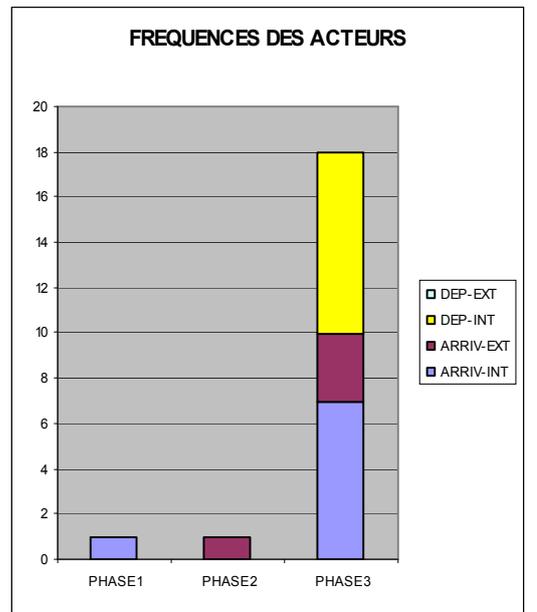
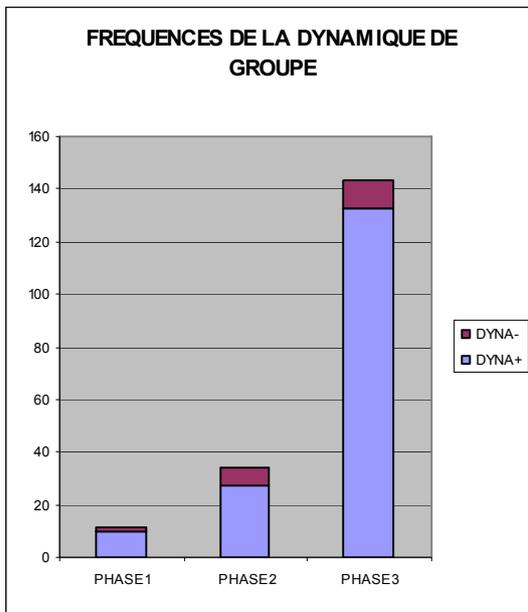
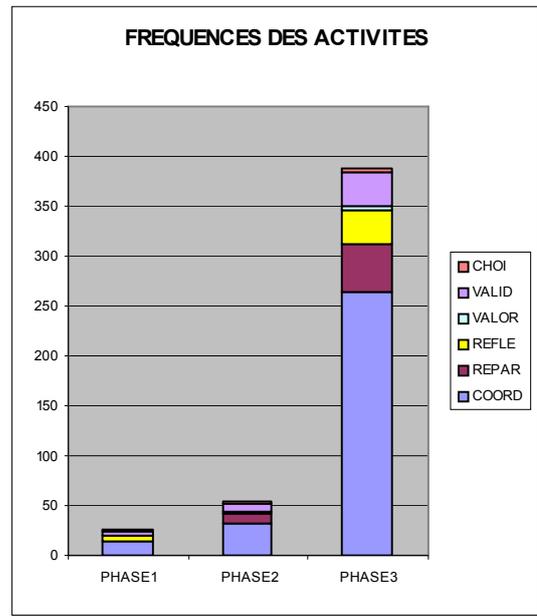
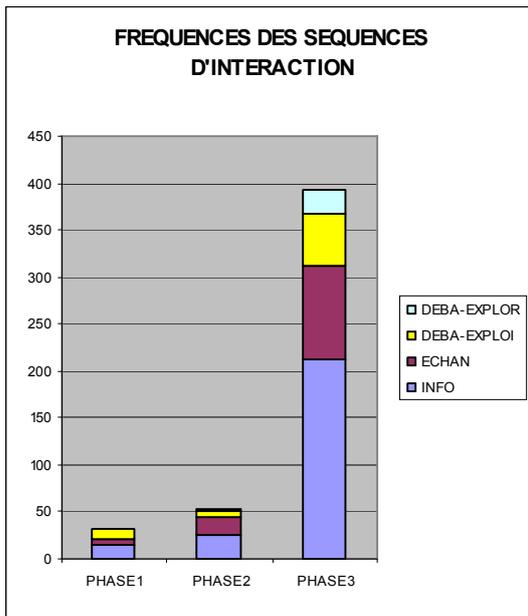
L'étape finale consiste, à partir de l'évolution de l'ensemble des trajectoires des différentes catégories de codes, à repérer les phases du processus de coopération. Nous précisons qu'une phase correspond au regroupement d'incidents dont les événements sont homogènes.

La construction des phases s'est appuyée sur la représentation graphique des trajectoires et sur une carte synthétique du processus de coopération telle que la propose Langley (1999). Cette matrice nous donne une vision synthétique de l'évolution de l'ensemble des trajectoires. La carte est représentée sous la forme d'un tableau présentant pour chacune des catégories de codes (en ligne) de manière chronologique (en colonne), les différentes périodes ayant émergé de l'analyse. Sur la carte, nous précisons par un code couleur les périodes définies au cours de l'analyse de chaque trajectoire.

Les phases sont représentées graphiquement en figure 3.

Nous avons rédigé un récit de synthèse de chaque cas de coopération présentant le processus de coopération et ses différentes phases.

Figure n°3 : Fréquences des codes des catégories dans les phases



1.2.5. PHASE 5 : INTERPRETATION ET VALIDATION DES RESULTATS

La méthode du codage des données a permis de faire émerger les phases du processus de coopération de manière plus précise et plus riche que les monographies et les sous-récits thématiques. En contrepartie, les récits effectués en première phase d'analyse (selon Yin, 1989) ont permis de mieux appréhender et comprendre le résultat du codage, ainsi que de l'enrichir.

Concernant la sous-question de recherche « Quelles sont les phases du processus de coopération ? », l'analyse issue du codage et des récits a fait émerger quatre phases de coopération : cadrage, répartition et définition des tâches, apprentissage de la coopération, apprentissage dans la coopération. La validation des résultats de la recherche a été obtenue à partir de la confrontation des résultats à la littérature existante ainsi qu'aux données recueillies.

2. PROPOSITIONS : APPORTS ET LIMITES DU CODAGE DES DONNEES

A partir de cette illustration du processus de codage, nous développons dans cette deuxième partie des propositions quant aux apports et aux limites du codage des données qualitatives.

2.1. QUELS APPORTS DU CODAGE DES DONNEES ?

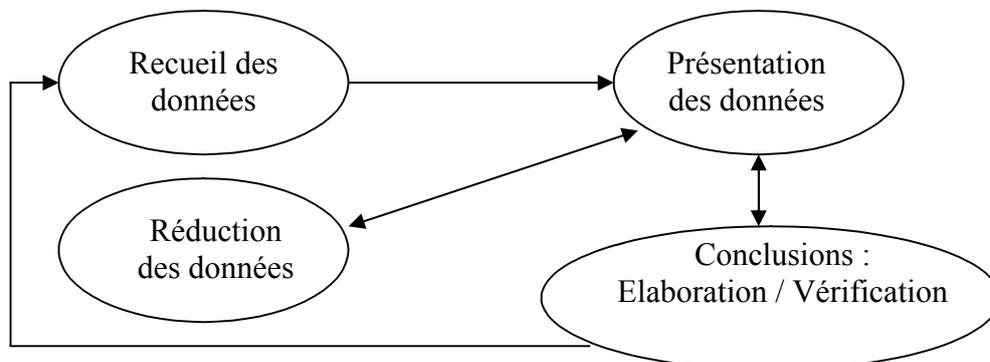
2.1.1. Le codage des données, un processus d'analyse des données interactif

Le codage de données qualitatives est un processus d'analyse des données complet.

Proposition 1 : *Le codage des données est un processus d'Analyse des données interactif.*

En effet, ce processus intervient dans chacune des composantes de l'Analyse des données proposées par Huberman et Miles (1991) : le recueil des données, la réduction des données, la présentation des données, l'élaboration /et la vérification des conclusions.

Figure n°4 : Composantes de l'Analyse des données : Modèle Interactif d'après Huberman & Miles (1991)



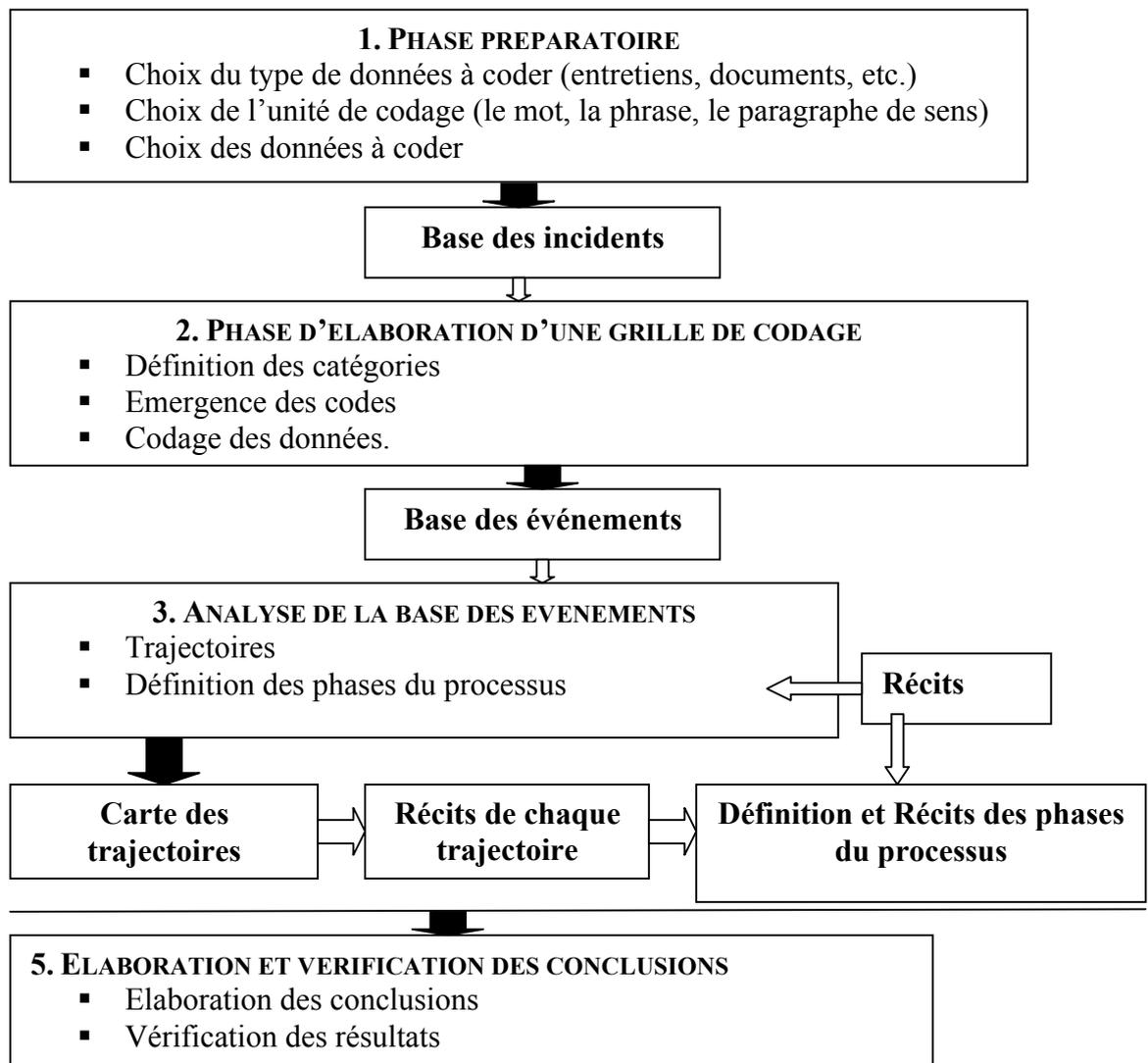
Dans le cas de notre recherche, le codage n'est pas intervenu dans la composante « Recueil des données » puisque nous sommes dans le cas d'un codage à posteriori. De la même manière que le processus d'analyse des données, le processus de codage est un processus interactif (cf. figure 5). De nombreux allers-retours sont effectués entre les différentes étapes. A partir de notre expérience du codage, nous considérons le codage comme un processus itératif en 5 étapes.

Proposition 1bis : Le codage des données est un processus itératif en 5 étapes :

- La préparation du codage des données
- La définition d'une grille de codage
- Le codage des données
- L'analyse de la base des événements
- Interprétation et validation des résultats

Chacune de ces phases est importante : les choix effectués par le chercheur au sein de chacune d'entre elles sont déterminants pour la qualité de l'analyse des données.

Figure 5 : le processus de codage des données qualitatives



Le processus de codage ainsi défini va dans le sens de Bardin (1977), qui aborde l'idée du codage comme un processus de transformation des données qui est réalisé selon des règles.

2.1.2. Le codage des données, un mode de présentation synthétique et systématique

Les codes définis précédemment nous ont permis de présenter les cas observés sous un même format afin de tirer des conclusions et d'établir des comparaisons.

Proposition 2 : *le codage des données est un mode de présentation synthétique et systématique des données qui favorise la définition des conclusions et la comparaison inter-cas.*

▪ **La comparaison inter-cas**

Le format de présentation systématique des données facilite la comparaison inter-cas et la crédibilise. Le codage des données permet le test des résultats de la recherche sur d'autres cas ultérieurs et donc leur généralisation. D'ailleurs, selon Miles & Huberman (1991, p.145) : « (...) dans le cadre d'une étude multi-site, si les formats de tous les sites sont semblables, le travail de l'analyse inter-site s'en verra considérablement facilité. »

▪ **Les résultats de la recherche**

La grille de codage permet d'analyser selon une trame de lecture commune et permet en même temps la prise de hauteur par rapport au terrain, la décontextualisation en ne conservant que les éléments de contexte déterminants.

Le codage des données nous a permis d'enrichir l'analyse des données et de définir des conclusions plus fines qui n'avaient pas émergé avec le récit initial. Cet enrichissement est lié à la manipulation importante des données. Le codage des données donne lieu à une synthèse des données qui permet une prise de recul plus importante. Ainsi, le codage favorise le détachement du chercheur au contexte. Par exemple, dans le cas de notre expérience, les phases du processus de coopération qui avaient émergé à partir de l'analyse du récit étaient très contextualisées. Le codage des données nous a permis de les transformer en phases générales décontextualisées permettant des comparaisons inter-cas.

Ainsi, dans le cas de notre recherche, si le récit nous a permis de définir les phases du processus de coopération, l'évolution des fréquences des codes dans le temps nous a permis de valider et de préciser ces phases par rapport à ces codes. La notion de phase d'apprentissage de la coopération est ainsi apparue. De même, l'idée de co-construction d'un outil est apparue avec le codage du contexte de création des outils.

Le codage permet la simplification d'un ensemble de données complexes. Cette méthode va au-delà d'un simple outil de mise en forme et de présentation des données, c'est un outil qui entame le processus d'analyse des données. En effet, selon Huberman et Miles (1991), cet assemblage organisé d'informations qui permet de tirer des conclusions et de passer à l'action. En effet, selon Huberman et Miles (1991) :

« Les codes (...). Ce sont des outils de recouvrement et d'organisation permettant à l'analyste d'identifier rapidement, d'extraire, puis de regrouper tous les segments liés à une question, une hypothèse, un concept ou un thème donnés. Ce regroupement ouvre la voie à l'analyse. » (Huberman & Miles, 1991, p. 94-97).

2.1.3. Le codage des données, un processus d'apprentissage des données

Les nombreuses manipulations auxquelles le chercheur doit se livrer lui donnent une connaissance et une maîtrise très importante de ses données, bien plus importante que celle qui résulte de l'écriture d'un récit. Ainsi, les nombreux allers-retours effectués pour définir une grille de codage enrichissent l'analyse des données.

Proposition 3 : *Le processus de codage des données est un processus d'apprentissage des données recueillies.*

Ces manipulations des données permettent la présentation des données et contribuent à l'émergence des résultats de la recherche. D'ailleurs, selon Mazars-Chapelon (2001, p.109) *« Un certain nombre d'outils inspirés des formats de présentation des données proposés par Miles et Huberman (1991). Il s'agit de matrices, figures et diagrammes. Certains constituent simplement des documents de travail. Ils permettront au lecteur d'accéder à des éléments de raisonnement qui ont été utilisés. (...) En revanche, certains formats de présentation constituent des étapes essentielles dans la conduite de l'analyse. »*

2.1.4. Le codage des données et l'évolution de la question de recherche

Le codage ne se résume pas à la constitution d'une grille de codage. Il s'agit d'un processus bien plus riche. En effet, le codage des données est un processus qui implique de nombreux choix pour le chercheur et qui suscite de nombreux questionnements.

Le codage fait évoluer la question de recherche. En effet, au travers d'une manipulation importante des données, il est donné au chercheur la possibilité d'entrevoir la richesse des données et les possibilités de résultats. Le codage est un processus de tâtonnement au cours duquel le chercheur affine sa question de recherche.

Proposition 4 : Le codage des données est un outil de co-construction de la question de recherche.

2.2. LES LIMITES DU CODAGE DES DONNEES : L'EVALUATION DE LA QUALITE DU CODAGE DES DONNEES

2.2.1. Une simplification de la richesse des données à outrance ?

Des chercheurs reprochent au codage la tendance de l'analyse des données à être centrée sur des chiffres au détriment du contexte, qui fonde la richesse des données qualitatives. Il est souvent reproché aux chiffres de bien vouloir dire ce que l'on veut. La crainte de voir le chercheur se contenter des résultats apportés par les résultats du codage et de prendre comme question de recherche la question correspondant aux résultats définis est en effet l'un des risques de la méthode.

Cela nous conduit à l'interrogation suivante : « le codage n'est-il qu'un moyen pour le chercheur de montrer ce qu'il peut ou ce que le codage veut bien montrer ? ». Le codage serait alors perçu comme un moyen de trouver des résultats de recherche en fonction des données, au détriment d'une question de recherche initiale.

A cette limite, nous répondons que la constitution d'une grille de codage et les nombreux allers-retours entre la grille et les données sont orientés par la question de recherche. La question de la qualité du codage et de son évaluation se pose alors.

2.2.2. La difficile évaluation des choix du chercheur : pertinence et lisibilité

Le codage des données n'est pas le gage d'une recherche de qualité. En effet, c'est davantage la qualité du processus de codage mené qui est importante. Au regard des nombreux choix réalisés par le chercheur pour définir les règles de codage des données, la pertinence de ces choix apparaît comme un gage de la qualité de la recherche tout aussi important. Il apparaît alors important que les travaux permettent de présenter ces choix afin que le lecteur puisse en évaluer la pertinence. Une meilleure lisibilité du processus de codage et des choix effectué par le chercheur est un moyen de contrôler davantage la qualité des recherches utilisant le codage des données.

2.2.3. Le codage, un processus coûteux en temps?

Si le codage des données conduit à une manipulation des données intensive, en contrepartie ce processus est très coûteux en temps.

2.2.4. Une double constitution de la grille de codage est-elle possible ?

La fiabilité du codage est évaluée par le double codage (Miles et Huberman, 1985). Pourtant, nous avons montré que la qualité du codage et des résultats était dépendante de la pertinence des choix effectués par le chercheur tout au long du processus de codage. Aussi, nous préconisons que la fiabilité du codage doit être contrôlée pour l'ensemble des phases préparatoires et de codage des données. Van de Ven & Poole (1990, 2002) ont effectué ce contrôle dans le cadre d'une étude sur les processus d'innovation parce qu'il s'agissait d'une recherche multi-chercheurs. Lorsqu'un chercheur conduit seul une recherche, il est difficile de trouver un chercheur qui puisse prendre le temps d'un tel travail de contrôle.

Le codage des données est alors adapté aux petites populations de cas.

2.2.5. La qualité du recueil des données, un gage de la qualité du codage

La qualité du codage effectué, ainsi que des résultats de la recherche, sont dépendantes de la qualité du recueil des données. Ainsi, la qualité du recueil des données est déterminante. En effet, les nombreux allers retours effectués entre la grille et les données pour la constitution d'une grille de codage implique la possibilité de faire évoluer la nature des données codées, les codes, voire les catégories. Cela implique que le recueil des données soit rigoureux et complet.

Proposition 5 : la qualité du codage et des résultats obtenus sont dépendants de la qualité du recueil des données.

3. CONCLUSIONS

Pour un nombre croissant de chercheurs, le codage des données est un signal de qualité d'une recherche. Nous avons montré que le codage des données est un véritable processus d'analyse des données au sens de Huberman & Miles (1991). Le codage est un moyen d'extraire la richesse de données qualitatives complexes et nombreuses.

La définition du processus de codage en cinq étapes est marqué de nombreux choix par le chercheur. La pertinence de ces choix est un élément important de la qualité menée. Le codage est un processus complexe, itératif, ce qui rend difficile sa description.

Il est important de retenir l'importance à accorder à la rigueur du recueil de données qualitatives du fait des éventuels changements de modalités de traitement ou des changements d'unités d'analyse, etc., et de son impact sur la qualité des résultats obtenus.

A notre question « le codage des données qualitatives est-il synonyme de qualité d'une recherche ? », nous répondons que la qualité du travail de codage est déterminante. Nous

ajoutons que la qualité d'une recherche qualitative est plus importante si l'on considère une voie alternative, prenant appui sur la richesse du codage des données et du récit.

Une troisième voie : le codage des données qualitatives et le récit, une forte complémentarité

Si la simplification de la lecture des données générée par le codage facilite la génération de propositions décontextualisées, le codage des données ne dispense pas d'une connaissance et de l'exploitation de la richesse des données des cas. Face aux nombreux allers retours entre le récit et le codage, qui se sont avérés indispensables, nous proposons une troisième voie, celle de la complémentarité entre l'approche classique du récit et de celle du codage. Notre position concernant le codage est que ce dernier ne doit pas se faire au détriment du récit, dont de nombreux auteurs ont montré les mérites (Pettigrew, 1992 ; Yin, 1989 ; etc.). Inversement, le codage apparaît comme un mode d'analyse complémentaire au récit auquel il apporte un éclairage important à son analyse. Le récit et le codage s'enrichissent mutuellement et permettent d'améliorer les résultats de la recherche, ainsi que la fiabilité et la validité de la recherche.

Cette voie alternative permet de dépasser le clivage classique entre approches qualitatives et approches quantitatives.

ANNEXE A - TABLEAU DE SYNTHÈSE DES CODES

CATEGORIE	CODE	DEFINITION
ACTEURS	ARRIV-INT	Arrivée interne
	DEPAR-INT	Départ interne
	ARRIV-EXT	Arrivée externe
	DEPAR-EXT	Départ externe
SEQUENCES D'INTERACTION	INFO	Information unilatérale
	ECHAN	Echange
	DEBA-EXPLOR	Débat de type Exploration
	DEBA-EXPLOI	Débat de Exploitation
ACTIVITE	REFL	Réflexion
	CHOIX	Choix
	REPAR	Définition et répartition des tâches
	COORD	Coordination
	VALID	Validation
	VALO	valorisation
EVENEMENTS DE LA VIE DES OUTILS	CREA	Création d'un outil
	UTIL	Utilisation d'un outil
	ADAPT	Adaptation d'un outil
	EVA	Evaluation d'un outil
	OUV	Ouverture d'un outil
	FERM	Fermeture d'un outil
	ABAN	Abandon d'un outil
OUTILS	GW	Groupe de travail
	DOC	Document
	TAB	Tableau
	AD	Aide à la décision
	CIRC	Circulation d'informations
	CAPI	Capitalisation
	PLAN	Plan de travail
	PHYSI	Outil physique

ANNEXE B - TABLEAU : LES CODES DE LA CATEGORIE « EVENEMENTS DE LA VIE DES OUTILS »

CODE	DEFINITION	COMMENTAIRES
CREA	Création d'un outil	La création d'un outil est un processus. Il s'agit de la période de temps qui précède la première utilisation de l'outil. En effet, la création d'un outil implique son élaboration lorsque celui-ci doit être construit, sa mise en place, etc. Cette étape peut comprendre l'évocation d'un outil qui n'existe pas encore.
UTIL	Utilisation d'un outil	L'utilisation d'un outil correspond à la mobilisation explicite d'un outil dans une intention d'action. Dans le cas d'outils ouverts, cela correspond également à une évolution de l'outil au niveau de son contenu .
ADAPT	Adaptation d'un outil	L'adaptation d'un outil correspond à la modification de la forme d'un outil existant.
EVA	Evaluation d'un outil	Lorsqu'un outil a été utilisé, l'évaluation correspond au moment où l'on parle de son fonctionnement en bien ou en mal de ses résultats.
FERM	Fermeture d'un outil	La fermeture signifie une formalisation, une cristallisation de l'outil qui ne pourra plus être interprété et modifié par les acteurs.
ABAND	Abandon d'un outil	Un outil est abandonné. Il ne sera plus utilisé par le groupe.

BIBLIOGRAPHIE

- ADLER, P., (1995), « Interdepartmental Interdependance and Cooperation : The case of the design / manufacturing interface », *Organization Science*, Vol.6, No.2, March-April, pp.147-167.
- BARDIN, L., (1977), *L'analyse de contenu*, Paris, Presses Universitaires de France (9ème édition : 1998).
- BARLEY, S.R., (1990), “ Images of imaging : notes on doing longitudinal field work ”, *Organization Science*, Volume 1, n°3.
- BELMONDO, C., (2002), La création de connaissances dans les groupes de travail, Thèse de doctorat, Université Paris-IX Dauphine, 453 pages.
- CECCHET, C., (1995), "Vers une méthodologie intégratrice pour l'analyse des problèmes de communication et de coordination", AIMS.
- CROZIER, M., FRIEDBERG, E., (1977), Le chercheur devant son terrain, in *L'acteur et le système*, ed. Seuil, pp.76-83.
- DAMERON, S., (2000), Génération de la coopération dans l'organisation - Le cas d'équipes projet, Thèse, Université Paris IX Dauphine.
- DAVID, A. (1998a), " Outils de gestion et dynamique du changement ", *Revue Française de Gestion*, Juillet.
- DAVID, A. (1998b), "Logique, épistémologie et méthodologie en sciences de gestion", Cahiers de recherche du DMSP, Juillet.
- EISENHARDT, K., (1989), "Building theories from case study research", *Academy of management Review*, vol.14, n°4, pp. 532-550.
- FORGUES, B., (1993), Processus de décision en situation de crise, Thèse de doctorat, Décembre.
- FORGUES, B., VANDANGEON-DERUMEZ, I., (1999), Analyses longitudinales, in R.A. Thiétart (ed), *Méthodes de recherche en management*, Dunod.
- HUBER, G. P., (1991), “ Organizational Learning : the contributing processes and the literatures ”, *Organization Science*, Vol. 2, No. 1, February, pp. 88-115.
- HUBERMAN, A. M., MILES, M.B., (1991), *Analyse des données qualitatives - Recueil de nouvelles méthodes*, De Boeck Université, 480 p.
- LANGLEY, A., (1999), " Strategies for theorizing from process data ", *Academy of Management Review*, 24 (4), pp .691-710.
- LEONARD-BARTON, D., (1990), “ A dual methodology for case studies : synergetic use of a longitudinal single site with replicate multiple sites ”, *Organization Science*, Vol 1, n°3, pp. 248-266.
- MAZARS-CHAPELON, A., (2001), “Outils de gestion, cognition, émotion, Thèse de Doctorat, Université Paris IX Dauphine.
- MORSE, J., (1994), “Designing funded qualitative research, in N.K. Denzin & Y.S. Lincoln (eds) *Handbook of Qualitative Research*, Thousand Oaks, CA, Sage, pp. 220-235.
- PELZ, D., (1985), “ Innovation Complexity and the Sequence of Innovating Stages ”, *Knowledge : Creation, Diffusion, Utilization*, Vol.6, No3, March, pp.261-291.
- PENTLAND, B. P., (1999), “ Building process theory with narrative : from description to explanation ”, *Academy of Management Review*, vol. 24, n°4, pp. 711-724.
- PETTIGREW, A.M., (1990), " Longitudinal field research on change : Theory and practice ", *Organization Science*, Vol.1; No.3, August, PP.267-292.
- PETTIGREW, A.M., (1992), " The Character and Significance of Strategy Process Research ", *Strategic Management Journal*, pp. 5-16.

- PETTIGREW, A.M., WOODMAN, R. W., CAMERON, K.S., (2001), " Studying organizational change and development : challenges for future research ", *Academy of Management Journal*, pp. 697-714.
- RING, P.S., VAN DE VEN, A.H., (1992), "Structuring cooperative relationships between organizations", *Strategic Management Journal*, Vol.13, n°7, pp.483-498.
- SAUBESTY VALLIER, C., (2004), « Codage de données qualitatives dans les processus de construction de coopérations en management de projets : le cas d'un "pôle confort" à la SNCF. », Séminaire "Les Ficelles du métier", ENS Cachan, Université de Marne la Vallée, Ecole Polytechnique, Ecole des Mines de Paris, Décembre.
- RING, P.S., VAN DE VEN, A.H., (1994), "Developmental Processes of Coopérative Interorganizational Relationships", *Academy of Management Review*, vol.19, No.1, pp.90-118.
- STRAUSS, J. CORBIN, A., (1998), Basics of qualitative research, 2d edition, Sage Publications.
- TERSSAC (DE), G., FRIEDBERG, E., (1996), Coopération et conception, Octares Editions.
- THIETART, R.A., (1999), Méthodes de recherche en management, Dunod, Paris.
- YIN, R. K., (1984), *Applications of Case Study Research*, Sage Publications.
- YIN, R. K., (1989), *Case Study Research : design and methods*, Sage Publications, seconde édition.
- VAN DE VEN, A.H., (1992), " Suggestions for studying strategy process : a research note ", *Strategic Management Journal*, vol.13, pp. 169-188.
- VAN DE VEN, A.H., HUBER, G.P., (1990), " Longitudinal field research methods for studying processes of organizational change ", *Organization Science*, vol.1, No.3, August, Special Issue, pp. 213-219.
- VAN DE VEN, A. H., POOLE, M. S., (1990), " Methods for studying innovation development in the Minnesota innovation research program ", *Organization Science*, Vol. 1, No. 3, pp.313-335.
- VAN DE VEN, A. H., POOLE, M. S., (1995), " Explaining Development and change in organizations ", *Academy of Management Review*, Vol.20, No. 3, pp.510-540.
- VAN DE VEN, A.H., POOLE, M.S., (2002), Field Research Methods, in J.A.C. Baum (ed.), *Companion to Organizations*, Oxford, Blackwell Publishers Ltd, pp.867-888.
- WACHEUX, F., (1996), *Méthodes qualitatives et recherche en gestion*, Economica, 1996.
- WEICK, K.E., (1995), *Sensemaking in Organizations*, Sage Publications.