



HAL
open science

Possibilités et pièges liés à l'utilisation des logiciels dans le processus d'analyse au travers de la théorie enracinée

Lionel Garreau, Rodrigo Bandeira-De-Mello

► To cite this version:

Lionel Garreau, Rodrigo Bandeira-De-Mello. Possibilités et pièges liés à l'utilisation des logiciels dans le processus d'analyse au travers de la théorie enracinée. Conférence AIMS, May 2008, Sophia Antipolis, France. halshs-00339851

HAL Id: halshs-00339851

<https://shs.hal.science/halshs-00339851>

Submitted on 19 Nov 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Possibilités et pièges liés à l'utilisation des logiciels dans le processus d'analyse au travers de la théorie enracinée.

**Rodrigo Bandeira-de-Mello, Université do Vale do Itajaí
(PPGAT/UNIVALI)**

**Lionel Garreau, Université Paris-Dauphine, DRM (CREPA), F-75016
Paris, France**

Correspondance à :

Lionel Garreau

6 passage Gauthier

+33 (0)6 72 97 08 77

lionelgarreau@hotmail.com

La méthodologie proposée par la théorie enracinée permet dans une optique idéale de proposer des résultats généralisables et objectifs. Néanmoins, de nombreuses critiques sont adressées aux recherches basées sur cette méthode. Les principaux reproches sont le manque de rigueur dans l'analyse des données et la mauvaise restitution du travail du chercheur dans sa recherche. L'utilisation de logiciels d'aide à l'analyse des données qualitatives permet de développer les possibilités d'analyse des données au sein de la théorie enracinée. Ces logiciels utilisent la puissance de l'outil informatique pour épauler le chercheur dans de nombreuses tâches nécessaires à l'élaboration d'une théorie enracinée de bonne qualité. En nous basant sur la littérature et une expérience de l'utilisation de plus de sept ans de ces outils, nous montrons de quelle façon ces logiciels permettent d'améliorer la qualité de la recherche basée sur la méthodologie enracinée. Nous mettons toutefois en garde les utilisateurs potentiels de tels logiciels en présentant cinq pièges auxquels peuvent être confrontés les chercheurs débutants comme les chercheurs confirmés. Nous espérons ainsi contribuer à la propagation de la bonne utilisation des logiciels d'aide à l'analyse des données, en rendant explicites les possibilités et inconvénients liés des ces logiciels.

Mots-clés : théorie enracinée, analyse des données, logiciel CAQDAS

Possibilités et pièges liés à l'utilisation des logiciels dans le processus d'analyse au travers de la théorie enracinée.

La méthodologie proposée par la théorie enracinée permet dans une optique idéale de proposer des résultats généralisables et objectifs. Néanmoins, de nombreuses critiques sont adressées aux recherches basées sur cette méthode. Les principaux reproches sont le manque de rigueur dans l'analyse des données et la mauvaise restitution du travail du chercheur dans sa recherche. L'utilisation de logiciels d'aide à l'analyse des données qualitatives permet de développer les possibilités d'analyse des données au sein de la théorie enracinée. Ces logiciels utilisent la puissance de l'outil informatique pour épauler le chercheur dans de nombreuses tâches nécessaires à l'élaboration d'une théorie enracinée de bonne qualité. En nous basant sur la littérature et une expérience de l'utilisation de plus de sept ans de ces outils, nous montrons de quelle façon ces logiciels permettent d'améliorer la qualité de la recherche basée sur la méthodologie enracinée. Nous mettons toutefois en garde les utilisateurs potentiels de tels logiciels en présentant cinq pièges auxquels peuvent être confrontés les chercheurs débutants comme les chercheurs confirmés. Nous espérons ainsi contribuer à la propagation de la bonne utilisation des logiciels d'aide à l'analyse des données, en rendant explicites les possibilités et inconvénients liés des ces logiciels.

Mots-clés : théorie enracinée, analyse des données, logiciel CAQDAS

INTRODUCTION

L'utilisation de logiciels dans une perspective de recherche a été développée en premier lieu pour l'analyse quantitative des données. La puissance de calcul des outils informatiques a permis aux chercheurs de gagner énormément de temps avec l'utilisation de tels logiciels. Aujourd'hui, de plus en plus de chercheurs utilisent l'outil informatique pour l'aide que ces outils peuvent apporter à une analyse qualitative des données. Depuis de nombreuses années, la recherche qualitative en sciences de gestion se heurte à une critique parfois latente, parfois plus explicite (Miles, 1979 ; Turner, 1983) de manque de rigueur. L'objectif de l'utilisation de ces logiciels est non seulement la production théorique, mais aussi l'augmentation de la qualité scientifique des résultats présentés. Pourtant, l'utilisation des logiciels dans la recherche qualitative ne garantit pas la rigueur et la qualité des résultats. Nous mettons ici en évidence les bénéfices et les pièges issus de leur utilisation. Le chercheur doit bien connaître les réelles potentialités du logiciel dans la démarche d'analyse choisie et éviter les pièges de son utilisation. La conscience conjointe de ces deux facettes est la condition nécessaire pour aboutir à des résultats de qualité.

Nous développerons notre discussion sur l'utilisation des logiciels en la délimitant dans le cadre de la théorie enracinée (Corbin et Strauss, 1990 ; Glaser et Strauss, 1967). Cette délimitation dans notre stratégie de recherche est nécessaire pour mieux montrer les bénéfices et pièges à partir d'exemples concrets. La constellation de stratégies de recherche qualitative, la diversité des fondements épistémologiques ainsi que la diversité des critères de rigueur et de qualité ne permettrait pas une analyse générale de la relation entre l'usage de logiciel et la qualité des résultats. Nous avons choisi la stratégie de la théorie enracinée pour trois motifs. Premièrement, la théorie enracinée est fréquemment évoquée comme une stratégie permettant d'améliorer la qualité de projets de recherches dont le but est de produire une théorie, et où le recours à la théorie enracinée se limite souvent à ses méthodes de traitement des données. Ensuite, l'emploi combiné de cette stratégie et des logiciels implique autant de bénéfices que d'écueils à éviter. Finalement, comme le soulignent Lonkila (1995) et Strübing (1997) la théorie enracinée est le contexte le plus adapté à l'utilisation des logiciels, car les deux mettent en valeur les activités de codage et le processus de recherche.

Ainsi, la combinaison de ces logiciels avec le recours de plus en plus fréquent à la théorie enracinée nous semble intéressante dans l'optique de produire des résultats scientifiquement

fondés, à évaluer selon des critères non positivistes (Corbin et Strauss, 1990). Dans cette optique, nous ne pouvons que nous réjouir de la référence de plus en plus fréquente à l'utilisation des techniques d'analyse de la théorie enracinée et à l'utilisation de logiciel comme appui au traitement des données, d'autant que ces logiciels et cette méthode permettent de tirer parti des données abondantes et accessibles à l'heure d'internet sur divers supports : photographies, bandes son, ou vidéo. Attention toutefois à ne pas tomber dans les pièges tendus à la fois par la théorie enracinée, dont les risques et problèmes inhérents peuvent devenir un cauchemar pour le chercheur (Goulding, 2001), et par l'utilisation du logiciel au cours de l'analyse des données.

Du reste, la référence fréquente à ces deux éléments (théorie enracinée et logiciel) peut être vue comme une recherche de légitimation de la part du chercheur, qui compterait sur la supposition que l'utilisation de ces techniques lui donnerait la crédibilité qu'il espère obtenir dans le champ académique (Lee et Fielding, 1996). Cette croyance tend à se réduire avec le nombre croissant de chercheurs qui connaissent, au moins partiellement, l'utilité de tels logiciels. Force est de constater que les recherches de qualité se fondent avant tout sur la réflexion du chercheur, et que les logiciels ne peuvent que l'aider dans cette tâche. A part quelques exceptions de manuels d'utilisation des logiciels spécifiques (Gibbs, 2002), la littérature traitant les méthodes de recherche qualitative n'adresse pas profondément les prescriptions sur « comment faire » l'analyse, en particulier avec l'aide de logiciels. On peut dire qu'il y a une sorte de tabou et de crainte que la formalisation puisse générer des mauvais résultats. Cet article remplit une partie de cet espace laissé vaquant en présentant une évaluation critique de l'usage des logiciels encadré par le processus d'analyse de la théorie enracinée. Pour ce faire, nous posons trois questions principales :

- a) Dans quelles mesures les logiciels peuvent-ils aider à la démarche proposée par la théorie enracinée ?
- b) Quelles sont les possibilités de l'utilisation des logiciels dans le processus d'analyse des données ?
- c) Quelles sont les pièges auxquels il faut faire attention ?

La discussion autour des avantages de l'utilisation des logiciels d'aide à l'analyse des données qualitatives n'est pas récente. Les méthodes présentées par Siedel et Clark (1984) pour THE ETHNOGRPAH, Richards et Richards (1991) et Richards (2002b) pour NUDIST et NVIVO, Muhr (1991) pour ATLAS.ti et Huber et Garcia (1991) pour AQUAD ainsi que l'évaluation générale de l'analyse de données qualitatives assistée par ordinateur de Bournois *et al.* (2002) montrent que l'usage des logiciels dans la recherche qualitative a depuis longtemps trouvé sa

place dans la communauté scientifique. Néanmoins, notre contribution avec cet article dépasse la présentation des fonctionnalités de quelques logiciels. Nous ne présentons pas ici un guide d'utilisation du logiciel ATLAS.ti ; le logiciel présente une grande flexibilité qui permet à chaque chercheur de produire sa propre démarche. Nous menons une discussion critique de l'utilisation de ce type de logiciels, spécifiquement dans le cas de la théorie enracinée, en révélant les bénéfices potentiels ainsi que les pièges.

Nous présenterons ainsi les bases de la théorie enracinée en mettant en avant les éléments qui poussent les chercheurs à utiliser cette méthode pour produire des résultats de recherche fondés scientifiquement. Nous verrons dans une seconde partie comment le logiciel peut fournir une aide précieuse pour chacune des étapes de la théorie enracinée, du recueil des données à la rédaction finale de la recherche. Enfin, nous exposerons les pièges auxquels peuvent être confrontés les chercheurs face à l'utilisation de logiciels d'analyse de données qualitatives.

1. LES FONDEMENTS CONCEPTUELS DE LA THEORIE ENRACINEE

La théorie enracinée, sur laquelle nous nous appuyons ici, n'est qu'une des multiples possibilités offertes au chercheur dans le domaine de la recherche qualitative. Développée par Glaser et Strauss en 1967, la théorie enracinée est « *une méthodologie d'analyse générale liée à la collecte des données, qui utilise un jeu de méthodes systématiques pour générer une théorie inductive sur une aire substantive* » (Glaser, 1992: 16). Elle présente certains avantages dans certaines situations comparativement à d'autres courants de recherche, mais ne s'impose en rien comme la meilleure méthode a priori à implémenter pour une recherche donnée ; elle peut en effet se révéler inadéquate au terrain, inadéquate au chercheur ou risquée dans sa mise en œuvre (Goulding, 2001). Il n'est pas question ici de faire une révision de la théorie enracinée ni de présenter l'exhaustivité des éléments qui composent cette méthode (Charmaz, 2000, 2006 ; Glaser et Strauss, 1967 ; Goulding, 2002 ; Locke, 2001 ; Strauss et Corbin, 1990), mais nous présentons ici les aspects du processus de la théorie enracinée pertinents pour notre analyse critique de l'usage des logiciels comme support au traitement de données au travers de cette démarche.

1.1. Conception et positionnement épistémologique de la théorie enracinée

La Théorie Enracinée a été développée par Barney Glaser et Anselm Strauss pendant les années 1960 comme une réponse aux attaques du courant dominant en sociologie dénonçant le manque de rigueur des recherches qualitatives. Glaser et Strauss avaient l'intention de

proposer une stratégie qui amoindrisse la subjectivité et l'intervention du chercheur, où les phénomènes étudiés seraient issus du terrain et non dérivés de la littérature. L'application de la théorie enracinée envisageait la recherche sur l'action d'un groupe spécifique au sein de situations encadrées (Goulding, 2002 ; Wells, 1995). Au cours des années, toutefois, les deux auteurs ont développé leurs propres versions de cette stratégie. La première, issue de l'ouvrage de Strauss et Corbin (1990), propose un système de codage complexe élaboré pour diriger le chercheur dans chaque étape de la recherche. La seconde (Glaser, 1992) propose une vision plus créative et reproche à son ancien collaborateur de forcer l'apparition de théories par un traitement mécanique des données. La méthode de Strauss et Corbin paraît moins risquée pour un chercheur inexpérimenté qui est guidé par à pas, et est moins tributaire du talent de création théorique du chercheur (Goulding, 2001). Néanmoins, les deux courants se rejoignent pour militer en faveur d'une certaine rigueur afin de proposer des théories robustes fondées sur des faits.

Les visions sur les fondements épistémologiques de la théorie enracinée divergent. Charmaz (2000), en critiquant l'inclination biaisée de la conception originale vers le positivisme, propose une vision de la théorie enracinée où l'implication du chercheur dans le recueil et le traitement des données tend à se rapprocher de la logique constructiviste. Glaser (2002) lui répond en avançant que la théorie enracinée ne peut pas être menée dans une perspective constructiviste. En effet, la logique même de la méthode est de parvenir à l'objectivité, comme le montre notre prochaine section, ce qui n'est pas le cas de la logique constructiviste. En revanche, l'implication de la sensibilité du chercheur et son influence sur le recueil et la teneur de certaines données sont communes aux deux courants (Glaser, 2002).

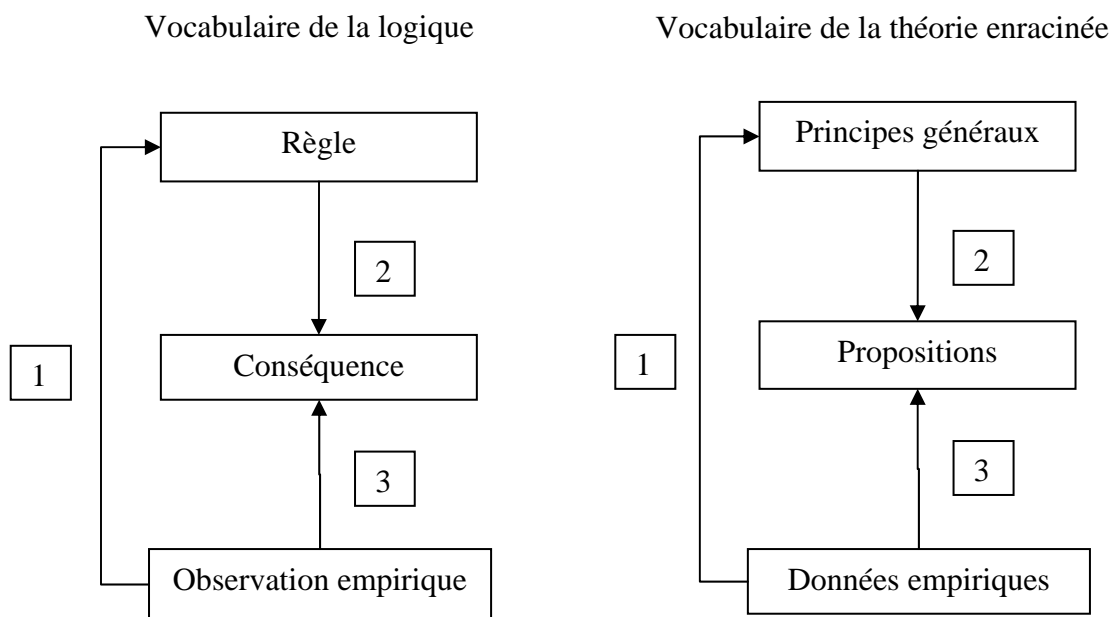
Néanmoins, les recherches menées sur le modèle de la théorie enracinée doivent respecter un certain nombre de critères précis, même si le chercheur doit trouver un équilibre entre une méthode stricte et les possibilités de réalisation de sa recherche (Corbin et Strauss, 1990). L'ensemble des concepts et catégories mises en évidence, avec la spécification de leurs liens doit ainsi permettre l'établissement d'une théorie enracinée dans les données issues du terrain. Nous pouvons donc placer la théorie enracinée à mi-chemin entre les extrêmes de l'objectivisme et du subjectivisme proposés par Morgan et Smircich (1980). Lorsque le premier assume une réalité externe où l'homme n'est qu'un agent réactif (posture du positivisme), le second considère la réalité comme formée par l'imagination humaine, en ayant la phénoménologie comme base épistémologique et stratégie de recherche. Corbin et Strauss (1990) voient dans la théorie enracinée une forte influence de l'interactionnisme et du

pragmatisme américain. Dans cette perspective l'homme est vu comme un acteur social, qui crée sa réalité par l'interaction symbolique. La réalité est subjectivée par l'interaction (ou l'inter-subjectivisme chez Berger et Luckmann, 1966).

1.2. Les logiques d'analyse au sein de la théorie enracinée : induction, déduction et abduction

Pierce conçoit l'objectivité comme la « fin idéale d'une processus récursif du type abduction – déduction – induction » (David, 1999). La force particulière de la théorie enracinée, où le chercheur fait évoluer de façon simultanée son construit théorique et son enquête sur le terrain, est de permettre au chercheur de développer une récursivité poussée de cycle décrit par Pierce. Le cheminement proposé par la théorie enracinée, en comparaison avec la logique proposée par David (1999), est présenté par la figure 1 ci-dessous.

Figure 1 : Description du raisonnement de la théorie enracinée ¹



La première étape de la théorie enracinée propose de trouver des principes généraux à partir des données recueillies sur le terrain. Cette étape correspond à une phase d'induction puisque nous proposons une règle générale qui devrait rendre compte de la situation étudiée par le chercheur. Dans un second temps, à partir d'un modèle général, le chercheur devrait pouvoir

¹ Le vocabulaire de la logique est emprunté à David, A., (1999), «Logique, épistémologie et méthodologie en sciences de gestion», *communication à Association Internationale de Management Stratégique*, Chatenay-Malabry

trouver des propositions logiques, conséquences directes des propriétés du modèle. Un raisonnement déductif est alors mis en œuvre. La troisième étape est fondamentale : il s'agit pour le chercheur d'induire des propositions logiques issues directement du terrain. Si ces propositions s'accordent avec les conséquences du modèle, alors c'est que le modèle est vérifié empiriquement par les faits, et donc valable dans les conditions observées. Le raisonnement est alors abductif car les données permettent de relier les conséquences du modèle à des faits réels. La théorie enracinée propose de passer par ces trois étapes, ce qui permet un contrôle plus fort de la qualité des recherches (Cicourcel, 2003), en révisant au besoin :

- Les règles quand elles ne permettent pas de rendre compte des éléments empiriques ou si leurs conséquences ne sont pas validées empiriquement ;
- Les conséquences si elles ne correspondent pas au modèle développé ou si les données empiriques ne s'accordent pas ;
- Les données empiriques si elles ne permettent pas de valider le modèle ou les propositions.

Ce dernier point, appelé échantillonnage théorique (Corbin et Strauss, 1990 ; Glaser, 1992 ; Glaser et Strauss, 1967), est particulièrement important. En effet, selon cette caractéristique de la méthode, au fur et à mesure de l'analyse et du recueil simultanés des données, le chercheur peut orienter sa recherche, en sélectionnant les incidents (ou l'occurrence du phénomène dans le matériau empirique) qui permettent au modèle et aux conséquences d'être vérifiés. L'échantillonnage théorique permet ainsi un retour aux données en fonction des orientations de la théorie en construction.

1.3 Qualité et rigueur pour la théorie enracinée

La logique d'analyse au sein de la théorie enracinée entraîne une façon particulière d'évaluer la recherche. D'abord, la qualité d'une théorie générée au sein de la théorie enracinée doit être mesurée par quatre critères proposée par Strauss et Corbin et synthétisé par Douglas (2003), que nous présentons dans le tableau ci-dessous.

Tableau 1 : Critères de qualité d'une recherche basée sur la théorie enracinée

Critères	Question
Adhérence	La théorie développée est-elle cohérente avec l'aire

	substantive étudiée ?
Intelligibilité	Est-ce que des personnes qui ne font pas partie de l'aire substantive étudiée peuvent comprendre la théorie élaborée ?
Possibilité de généralisation	Est-ce que la théorie développée s'applique à un nombre élevé de situations dans l'aire substantive étudiée ?
Contrôle	Est-ce que la théorie permet à l'utilisateur un certain degré de contrôle sur les structures et les processus des situations quotidiennes qui évoluent dans le temps ?

Adapté de Douglas (2003)

Au vu de ces critères de qualité et à l'utilisation de l'échantillonnage théorique, le chercheur devra témoigner de l'ensemble de son processus de recherche et d'analyse lors de la restitution des résultats afin de ne pas donner une image biaisée de la recherche, en particulier afin de ne pas faire croire que l'ensemble des cas potentiels correspondent au modèle développé. Le chercheur doit alors expliciter les raisons de ses choix de sélection de terrain(s) de recherche et donc quel contexte (ou « aire substantive ») prévaut dans l'établissement de son modèle. Un travail supplémentaire peut alors avoir lieu pour analyser les cas qui ne correspondent pas au modèle développé : soit que l'aire substantive ne correspond pas aux limites émises dans le modèle, soit qu'il existe des variantes qu'il reste à découvrir pour rendre compte de l'ensemble des situations possibles au sein de l'aire substantive étudiée.

La situation décrite ci-dessus apparaîtrait absurde dans une logique positiviste, alors qu'elle est naturelle dans l'optique de la théorie enracinée. Ici il n'est pas question pour la recherche de remplir les critères de fiabilité et de validité propres aux études sous une logique positiviste. Néanmoins, pour mieux comprendre la relation entre l'usage des logiciels et la qualité et la rigueur de la recherche, nous proposons l'interprétation des canons de la rigueur positiviste dans le cadre de la théorie enracinée.

Tableau 2 : La rigueur pour la théorie enracinée

Canons	Interprétation positiviste	Réinterprétation pour la théorie enracinée
Validité Interne	Mesure de la cohérence entre les éléments de la théorie vérifiée, ses	Étant donné que la réalité est socialement construite à travers la perception des individus, le chercheur doit identifier cette réalité dans

	interrelations, et la réalité.	les matériaux empiriques pour bien développer ses propositions.
Fiabilité	Mesure de la probabilité que la recherche, si répliquée, produise les mêmes résultats.	Il s'agit de la consistance ou cohérence du processus de recherche. Il s'agit donc ici de la fiabilité du processus de recherche, et non des résultats. Un autre chercheur, en suivant le même processus de recherche, peut avoir les mêmes résultats. Mais il n'y a pas un seul processus possible : deux chercheurs peuvent arriver à des résultats différents à partir des mêmes matériaux.
Validité Externe	Mesure de la possibilité de généralisation des résultats dans d'autres contextes comme dans l'inférence statistique.	Le but d'une théorie développée au sein de la théorie enracinée est d'expliquer comment et pourquoi des différentes manifestations du phénomène en question se produisent dans le cadre de l'aire substantive. Donc, plus l'échantillon est varié, plus la variabilité est expliquée par la théorie, qui devient ainsi plus générale.

Adapté de Merriam (1998) et Strauss et Corbin (1998)

2. POSSIBILITES LIEES A L'UTILISATION DES LOGICIELS POUR L'ANALYSE DES DONNEES QUALITATIVES AU SEIN DE LA THEORIE ENRACINEE

Loin de faire l'unanimité dans la communauté scientifique, les logiciels pour l'analyse des données qualitatives apportent autant de possibilités que de pièges. Ces logiciels présentent des limites auxquelles le chercheur doit porter attention. Ces limites latentes et souvent méconnues rendent dangereuse l'utilisation de la technologie d'information comme les logiciels d'aide à la recherche qualitative. Les possibilités sont elles aussi obscures dans la mesure où l'on ne comprend pas les vraies fonctionnalités et les contributions du logiciel dans le processus d'analyse.

Dans les sections suivantes, nous identifions les possibilités issues de l'utilisation d'une catégorie spécifique de logiciels proposant une aide au chercheur, appelés CAQDAS

(*Computer-Aided Qualitative Data Analysis Software*²), dans la démarche d'analyse proposée par la théorie enracinée de Strauss et Corbin (Strauss et Corbin, 1990). Cette catégorie de logiciels diffère de ceux dont la fonction principale est le comptage de mots, la fréquence de l'occurrence des expressions dans le texte, ou encore le simple codage de données avec un dictionnaire de thèmes préalables. Les logiciels qui appartiennent à cette catégorie, les CAQDAS, aident à la découverte de catégories et d'une structure latente dans les données. Ils aident le chercheur à organiser et enregistrer les processus d'analyse ainsi qu'à communiquer les résultats produits (Weitzman et Miles, 1995). La discussion suivante a pour but de montrer la relation entre l'usage de ces logiciels CAQDAS et la qualité et la rigueur de la théorie enracinée, à partir des récits issus de la littérature et sur l'expérience des auteurs de plus de sept ans d'usage du logiciel ATLAS.ti en tant que doctorant, chercheur et directeur de recherche.

L'utilisation de ce type de logiciel dans la recherche qualitative a été « popularisée » dans les années 80 par les premières versions de THE ETHNOGRAPH. Cependant, l'introduction de la technologie d'information dans le traitement qualitatif de données a connu des débuts difficiles à cause du scepticisme des chercheurs et de la peur d'une mécanisation de la recherche qualitative. Il est évident aujourd'hui que l'interface d'alors, MS-DOS, présentait des inconvénients forts. L'alternative qui restait la plus populaire était donc de couper les morceaux des transcriptions et les archiver dans les enveloppes, chacune désignant une catégorie analytique. Malgré les résistances initiales, des nouveaux logiciels et leurs améliorations au travers des versions successives ont été lancés, stimulés par l'environnement graphique des fenêtres qui a beaucoup apporté pour atténuer les inconvénients de travailler directement sur l'écran d'ordinateur. Les logiciels ont donc trouvé leur place dans la recherche qualitative (Lee et Esterhuizen, 2000).

2.1. Des fonctionnalités principales

Les fonctionnalités des logiciels CAQDAS sont développées dans l'optique générale d'organiser et d'enregistrer les processus d'analyse et aussi d'aider à communiquer les résultats. Il ne s'agit pas de l'analyse de contenu quantitative basée sur la fréquence de mots, expressions et concordances, soumise au traitement statistique, mais, plutôt l'analyse herméneutique où le chercheur essaye de découvrir des éléments de sens dans le texte ou d'autres supports comme les images. Il faut d'abord mettre en évidence que le logiciel aide

² Bournois et al. (2002) utilisent la version française de cet anagramme qui prend la forme ADQAO (Analyse de données qualitative assistée par ordinateur)

seulement le chercheur en organisant les données et en registrant la démarche de l'analyse. Il ne traite rien, il n'analyse rien. La compréhension du texte et de la structure latente dans les données, dans l'esprit du *Verstehen* (Abel, 1948 ; Moran, 2002) appartient toujours au chercheur (Kelle, 2002). Par conséquent, l'expression « logiciels d'analyse de données qualitatives » peut générer une certaine confusion. Nous suggérons plutôt la désignation « logiciels de support à l'analyse qualitative de données ».

Si les logiciels n'analysent pas, que font-ils alors ? Ils aident au classement, à l'organisation des éléments de la recherche et à la reconstitution du processus d'enquête, ce qui contribue à l'améliorer la qualité et la rigueur de la recherche. Le tableau 3 présente les fonctionnalités principales des logiciels CAQDAS du type générateur de théorie basée sur des codes. Si les fonctionnalités peuvent varier, ce tableau tend à rendre compte des caractéristiques communes aux principaux logiciels CAQDAS utilisés pour la recherche en gestion.

Tableau 3 – Fonctionnalités principales des logiciels CAQDAS

Fonction	Description
Sauvegarder et gérer les données	Le logiciel sauvegarde d'une façon intégrée, dans le cadre d'un projet, et indexée selon les critères du chercheur, toute la base de données, qu'il s'agisse de textes, sons, images et vidéos. Il gère aussi la fusion de différents projets.
Coder et associer	Le processus de codage permet d'associer un morceau de données, soit une partie de texte, une section d'une image ou un segment de vidéo/audio, à un code (ou, dans le langage de plusieurs stratégies de recherche qualitative, une catégorie qui représente une idée, un concept). L'association entre codes est aussi possible. Le codage ainsi que l'association donnée / code sont définis par le chercheur et non pas par le logiciel.
Chercher, saisir et récupérer des informations dans leur contexte	Les « outils de recherche » aident le chercheur de plusieurs manières : ils cherchent des mots et expressions, ils aident à identifier les occurrences d'un certain code dans les données, ainsi que la combinaison de codes selon certains critères tels que ceux de proximité ou de cooccurrence. Les résultats sont fournis dans leur contexte, c'est-à-dire les alentours de la section de données récupérées (texte, l'image, audio ou vidéo).
Contribuer dans	Le chercheur peut enregistrer ses <i>insights</i> et aussi les décisions prises au

le développement d'un modèle ou d'une théorie.	cours de la démarche d'analyse d'une façon à pouvoir la reconstituer et la restituer ultérieurement. L'interface graphique aide à atténuer l'intangibilité du processus dans la mesure où les associations entre les codes (ou d'autres éléments) sont visible au fur et à mesure qu'elles sont définies par le chercheur.
Aider à l'évaluation de la qualité de la recherche	En plus de pouvoir représenter graphiquement ses interprétations, le chercheur peut compter sur plusieurs rapports qui lui fournissent, par exemple, la chronologie du travail analytique, la hiérarchie des codes associés et leurs occurrences dans le texte. Les évaluateurs du travail peuvent analyser ces documents pour juger de la qualité de la recherche selon les critères de la stratégie employée.

Source : Adapté d'Alexa et Zuell (1999), Weitzman (2000) et Lacey et Luff (2001)

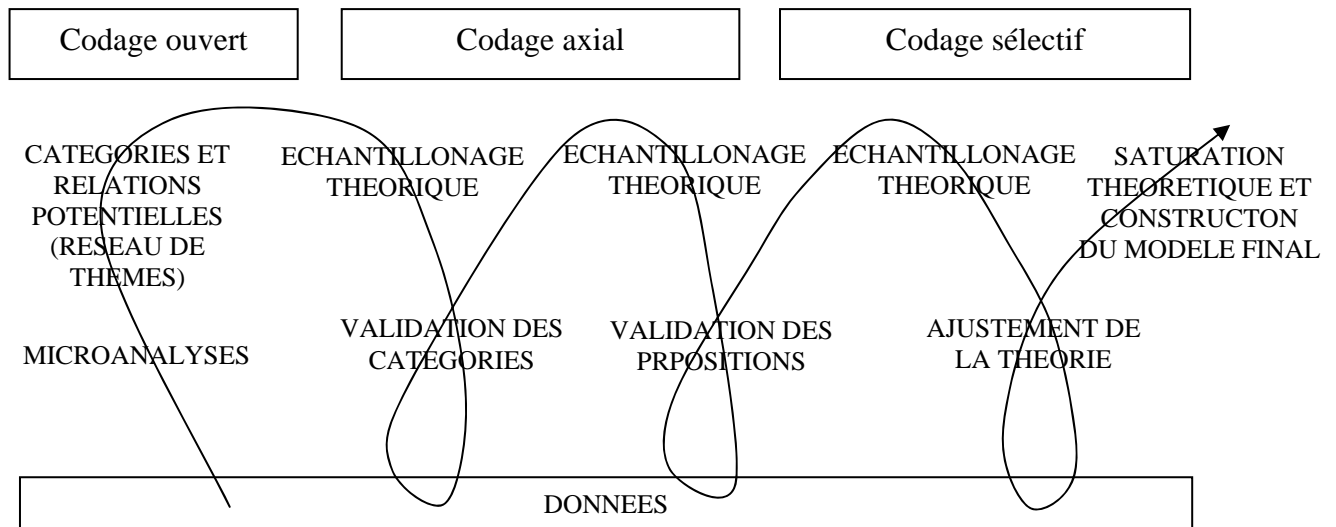
Les fonctionnalités principales décrites dans le tableau 2, si elles sont bien exploitées, peuvent aider le chercheur dans le processus d'analyse proposé par la théorie enracinée. Pourtant, le logiciel, même avec l'ensemble de ces fonctionnalités, ne propose pas d'analyse. C'est au chercheur de comprendre les fonctionnalités et les outils du logiciel, de pratiquer pour bien connaître et utiliser le potentiel offert par le logiciel, puis d'estimer de quelle façon le logiciel peut aider dans la conduite de la recherche. L'une des qualités principales du logiciel ATLAS.ti est sa grande flexibilité d'utilisation. Notre travail ici ne sera pas de faire adopter une démarche formatée aux chercheurs à travers ce logiciel, mais de montrer comment ATLAS.ti peut permettre de proposer des analyses et donc des résultats qui correspondent aux critères de qualité établis par la théorie enracinée.

2.2. Le processus d'analyse à l'aide du logiciel

La production d'une théorie enracinée implique le passage par de nombreuses étapes qui permettent au chercheur d'aboutir à des conclusions objectives et généralisables sur l'occurrence d'un phénomène dans une aire substantive. L'objectif ici n'est pas de passer en revue l'ensemble de ces étapes, mais d'offrir une vision globale de ce processus afin de comprendre en quoi et à quel moment le logiciel CAQDAS pourra aider à la production d'une théorie de qualité. La figure ci-dessous présente ainsi le processus parcouru par le chercheur dans l'optique de la création d'une théorie enracinée. Elle représente deux niveaux : au niveau « données », on trouve les discours des sujets, et tous les autres matériaux empiriques. Au dessus, un niveau « conceptuel » où les trois codages sont développés par le chercheur. Le

voyage entre ces deux niveaux caractérise l’aller-retour entre découverte et validation, typiques de la logique abductive.

Figure 2 : processus de la théorie enracinée



2.3. Recherche des éléments significatifs dans les matériaux : le codage ouvert

Deux approches se distinguent dans l’application de la théorie enracinée : la première, issue des travaux de Glaser, Strauss et Corbin, qui est la stricte application de la théorie en vue de développer une nouvelle théorie. Un second courant se démarque de la version « puriste » de la théorie enracinée, et regroupe les recherches qui font appel aux méthodes d’analyse de la théorie enracinée en les appliquant au sein d’autres stratégies de recherche. Ce second courant identifie en amont de l’étude terrain les principaux éléments à étudier. Cependant, cette vision pose le problème de forcer les données (Kelle, 2005) afin de les faire s’adapter à des éléments théoriques préexistants. Si cette option semble aller à l’encontre des prémisses de la théorie enracinée, elle permet néanmoins de limiter le champ d’étude, de cadrer le chercheur dans la pléthore de données issue de l’étude empirique, et de contribuer à l’amélioration des modèles existants dans une logique d’accumulation de la science vers une optique de généralisation des résultats (Kaufmann, 2007). Notre recherche sur l’utilisation combinée des logiciels CAQDAS et de la théorie enracinée se place dans le premier courant, approche où l’étape de codage ouvert est la plus importante dans l’émergence de nouvelles théories.

Pendant la phase de codification ouverte, étape initiale du processus, les activités de microanalyse et de codage sont intensives (Corbin et Strauss, 1990 ; Goulding, 2001 ; Strauss et Corbin, 1990). La production de microanalyses consiste en la déconstruction d'éléments des données : mots, groupes de mots ou phrases lorsqu'il s'agit de données textuelles. L'objectif est de comprendre ce qui est contenu dans les propos de l'acteur mais qui n'est pas rapporté de façon explicite afin de faire émerger des catégories potentielles et d'augmenter le nombre de voies d'interprétations possibles (Strauss et Corbin, 1990). Les métaphores ou les analogies dans les données textuelles se prêtent particulièrement bien à ce type d'exercice. Cette étape d'analyse permet de faire le lien entre verbatim et codes potentiels. Elle assure ainsi une meilleure qualité d'adhérence de la théorie proposée aux données empiriques. L'encadré 1 illustre ce type d'analyse, étape centrale dans la production de théories de bonne qualité (Goulding, 2001). Néanmoins, cette étape difficile est tributaire de la sensibilité théorique du chercheur (Glaser, 1978).

Encadré 1 : extrait d'une micro analyse

Citation du chef de projet :

« Voilà un peu ce que l'on peut en dire, alors...donc la particularité du dossier, c'est comme je te disais, euh, on nous a euh,...dans le *panier de la mariée*, il y avait aussi une commercialisatrice [...] »

Qu'est-ce que la métaphore du panier de la mariée veut dire au-delà de l'utilisation courante de l'expression ?

Aux yeux du chef de projet, un projet c'est un "Panier". Alors, quelles sont les propriétés d'un panier?

1. Ils gardent, ou bien, sont composés de l'ensemble des objets de même nature ou de natures différentes. C'est un "package".
2. Ils sont utiles puisque ils facilitent le transport et la logistique nécessaires pour gérer cet ensemble.
3. Par contre, il faut qu'on achète toute la diversité contenue dans le panier, pas une seule chose. C'est un "lot".
4. La concentration des choses dans un même panier ne diversifie pas le risque de perte si le panier est cassé ou perdu ou volé.

Ces caractéristiques et propriétés d'un panier peuvent être importantes pour comprendre comment les différents types de projets influencent-ils le travail du chef de projet.

On peut formuler des hypothèses selon lesquelles les caractéristiques d'un panier / d'un projet influencent le travail du chef de projet:

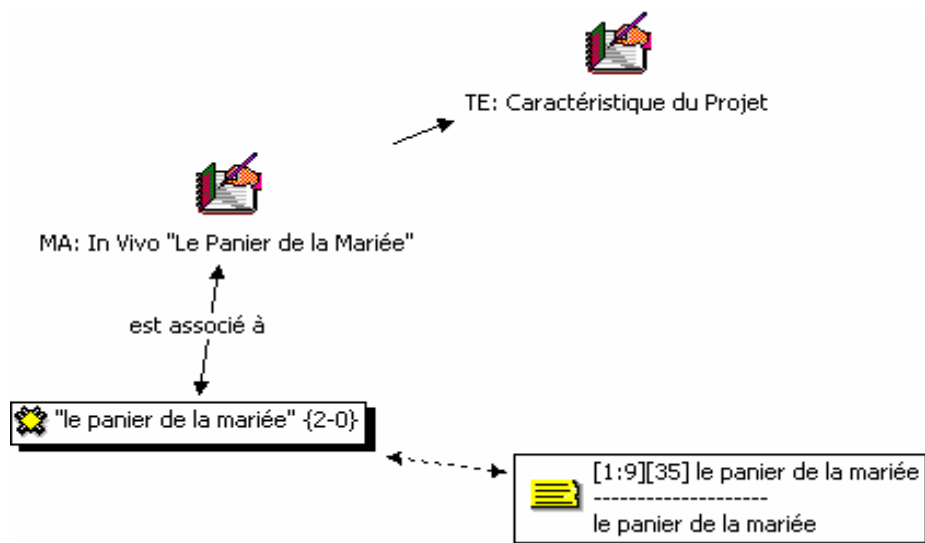
1. Ce projet sort du cadre habituel de l'organisation, puisqu'il mêle commerce et loisirs alors que d'ordinaire l'entreprise ne fait que du commerce. Le chef de projet doit gérer cet ensemble avec une diversité qu'il ne maîtrise pas bien.
2. La distribution imposée des rôles des parties du partenariat diminue l'impression de contrôle par le chef de projet. C'est à dire que le chef de projet doit gérer un ensemble qu'il n'a pas choisi lui-même.
3. On trouve la diversité du projet lors des réunions où chaque partie prenante défend ses propres intérêts. Donc, plus la nature des parties prenantes est hétérogène, plus on a une diversité dans le projet. Pour éviter tautologies, il faut définir quelles sont les conditions qui entraînent la diversité. Sur quels critères se base cette diversité ?

[...]

La microanalyse génère un grand nombre d'informations, dont la gestion est fondamentale pour que le chercheur ne se perde pas (Glaser, 1978). Pour cela, les fonctions de sauvegarde des données dans une base organisée afin de récupérer les informations ultérieurement sont utiles. Une fois les chercheurs libérés des tâches mécaniques d'enregistrement, indexation et récupération, il leur est possible d'approfondir la réflexion et l'interprétation face aux données. Cette optimisation aide à produire une théorie de qualité dans un délai potentiellement restreint grâce au logiciel. En outre, la possibilité de générer des présentations visuelles des interprétations est aussi utile pour que la recherche réponde aux critères de validité interne et fiabilité.

Dans le cas d'ATLAS.ti, les « mémos » peuvent être utilisés pour l'enregistrement de tout ce que le chercheur veut, y compris ses interprétations. Les « mémos » peuvent aussi être liés les uns avec les autres pour former un réseau logique d'idées, lequel peut représenter une construction primaire de catégories et ses rapprochements. Les codes qui font référence aux idées désignées par les memos peuvent être liés aux « memos » respectifs, ce qui donne un premier ordre aux codes créés. Les codes qui font référence aux verbatim du texte peuvent être traités de façon assistée grâce à l'aide du logiciel. Le codage par liste permettra de reprendre les codes déjà créés par le chercheur et toujours disponibles, le codage ouvert de créer un nouveau code face à un verbatim relevé dans le texte, et le codage *in vivo* de désigner la citation codée par un code du même nom. La figure 3 présente réseau de thèmes en construction avec l'association d'une citation, d'un codage *in vivo*, d'un mémo composé d'une microanalyse ainsi que d'un thème.

Figure 3 : Constitution d'un réseau de thèmes



Tous ces éléments peuvent être récupérés, c'est-à-dire qu'en cliquant sur l'objet, le chercheur sait ce qu'il signifie et où sont ces occurrences dans les matériaux empiriques. La familiarité du chercheur avec l'ensemble de données brutes et des éléments créés (codes, memos, citations) est particulièrement importante pour arriver à un haut niveau de validité interne.

La méthode de comparaison constante permet de mettre en lumière les aspects significatifs des situations décrites. Cette méthode permet de dépasser le caractère d'unicité de chaque situation en comparant différentes situations les unes aux autres, et en multipliant les options pour définir les catégories pertinentes pour la recherche (Friedberg, 1994). Ainsi, les éléments pertinents peuvent être choisis ou découverts grâce à une systématisation de la comparaison entre deux situations, que l'accent soit mis sur les caractéristiques communes des incidents traités, ou que ce soit les différences qui interpellent le chercheur. Les analyses alors produites permettent au chercheur de trouver les directions dans lesquelles il pourra conduire sa recherche, en identifiant les concepts sur lesquels il pourra produire son analyse (Goulding, 2002). Le logiciel permet alors de commencer à organiser les codes entre eux ainsi que d'éviter les redondances de codes en les classant dans des familles. Ces familles de codes permettent alors l'organisation des idées du chercheur autour de thèmes qui émergent. L'ensemble de ces interprétations et liens établis donne des indications qui guident le chercheur vers la formulation des associations probables, base du modèle ou de la théorie, au cœur du codage axial.

2.4. De la relation entre les éléments pertinents : le codage axial

Le codage axial consiste en l'établissement d'un niveau d'abstraction supérieur grâce à la spécification des liens entre des concepts à travers leurs propriétés. C'est le moment où certaines propositions, traitant des relations entre certains éléments du terrain, peuvent être formulées afin d'être validées ultérieurement. Ainsi, en suivant l'inférence abductive omniprésente dans le codage axial, le chercheur fonde et valide empiriquement les associations. Dans cette démarche, les codes sont la « pierre de touche » de la théorie, c'est-à-dire que le tri entre les éléments et / ou associations significatifs et non significatifs est à la base du développement des modèles et théories.

Dans cette phase, les logiciels CAQDAS peuvent aider les chercheurs de plusieurs manières. L'interface graphique, présentant ces codes permet de représenter en temps réel, et au fur et à mesure, les associations qui forment le modèle ou la théorie en construction. La récupération

automatique des informations résultant des recherches sur l'occurrence de codes dans le texte fournit au chercheur le contexte où les codes ont été identifiés. Ces matériaux empiriques servent ainsi à l'élaboration des propositions provisoires : le contexte dans lequel les codes se manifestent peut fournir des conditions, des conséquences et des stratégies d'action (Strauss et Corbin, 1990), et de nombreux indices afin d'établir de nouveaux rapprochements et de nouveaux codes en suivant ainsi la caractéristique circulaire et non-linéaire des différentes étapes du codage (figure 2) pour expliquer l'occurrence du phénomène identifié. Ces outils de recherche rendent possible une « attaque des données » dans la mesure où il est possible de combiner les critères de recherche et d'explorer les données selon de nombreux critères choisis par le chercheur. Au delà de l'amélioration des critères de validité interne et de fiabilité, la fonctionnalité d'exploration organisée de plusieurs contextes permet d'accroître la généralisation des résultats (validité externe) dans la mesure où l'aire substantive est agrandie sans perte de qualité d'analyse.

Dans le cas d'ATLAS.ti, l'association de type transitif désigne une liaison établie entre deux codes, dont, un code supérieur, abstrait (sans association directe aux données) se manifeste dans les données à travers les codes inférieurs, empiriques (liés directement aux données). Le niveau d'abstraction est donc plus élevé dans le cas du code supérieur, mais l'établissement de la hiérarchie entre les codes permet de travailler au niveau conceptuel et d'entrer dans la construction théorique en préservant le lien avec le niveau des données. Le classement de codes contribue grandement à l'amélioration de la qualité de la recherche qualitative dont le but est la création ou l'amélioration théorique. L'aide graphique apportée par le logiciel est indéniable, d'autant plus que les relations spécifiées peuvent être revues et corrigées au fur et à mesure qu'elles sont testées empiriquement sur les données. Le logiciel permet donc une grande flexibilité dans l'organisation des relations entre codes ainsi qu'un outil de traçage efficace afin de comprendre et restituer le processus de la recherche menée. Cela permet au chercheur de travailler avec des concepts plus généraux sans perdre le contact avec la réalité du terrain. La formalisation du processus rend tangibles les associations effectuées, ce qui permet à d'autres chercheurs de le vérifier et d'améliorer la fiabilité (ATLAS.ti gère aussi le travail de plusieurs chercheurs au sein d'un même projet).

Ici encore, les comparaisons doivent être effectuées de manière systématique afin d'affiner le codage axial. Kaufmann (2007) avance que c'est en comparant systématiquement les occurrences des événements ou des thèmes dans le discours des personnes interviewées que l'affinage des relations entre les concepts se réalise. Le test des codes axiaux sur l'ensemble des données permet, s'il est positif, de valider le critère de possibilité de généralisation des

résultats. En outre, les exceptions à une récurrence identifiée permettent d'affiner la théorie produite en formulant des hypothèses quand aux conditions de réalisation des phénomènes. Cela permettra dans la troisième étape de sélectionner les catégories centrales des catégories périphériques.

2.5. La construction du noyau théorique : le codage sélectif

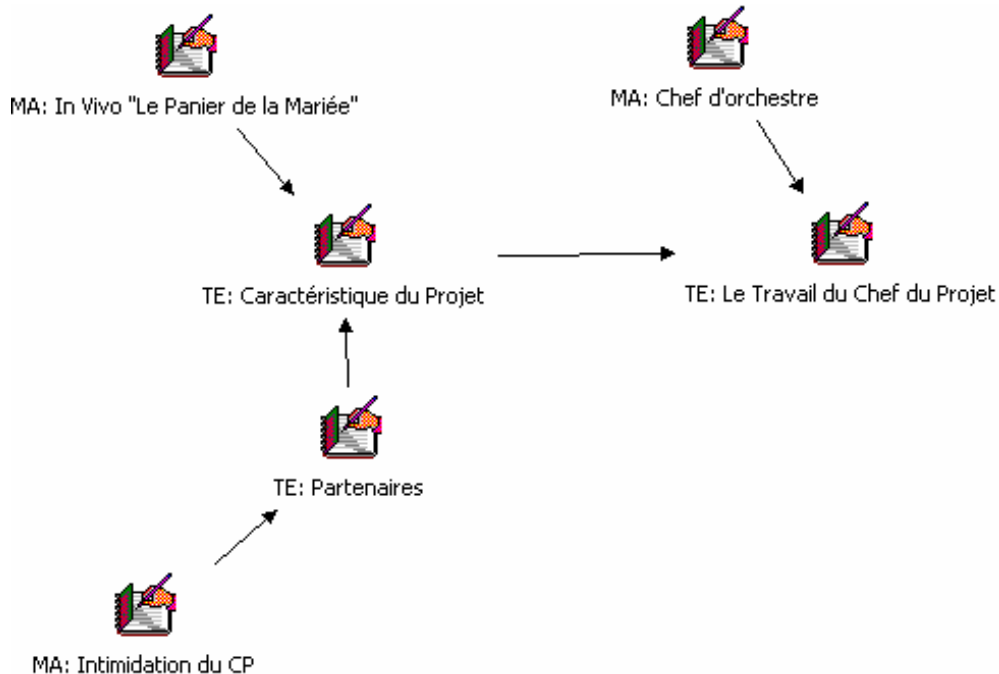
Une fois les éléments importants spécifiés et les relations entre les concepts établies, une troisième étape est nécessaire afin de produire une théorie enracinée qui réponde aux critères de qualité d'une telle recherche : l'établissement d'une catégorie centrale (Corbin et Strauss, 1990). Le codage sélectif est le processus qui permet au chercheur de déterminer la catégorie centrale autour de laquelle la construction théorique pourra prendre forme. Strauss et Corbin (1990: 14) proposent de se poser les questions suivantes afin de pouvoir identifier la catégorie centrale : « *Quelle est l'idée analytique principale présentée dans cette recherche ? Si mes recherches devaient être synthétisées en quelques phrases, qu'est-ce que je dirais ?* ». Cette phase du travail nécessite à la fois des compétences en terme de capacités d'abstraction de la part du chercheur, mais aussi de parcimonie car il peut avoir à choisir entre plusieurs catégories qui semblent égales sur un niveau de centralité.

Le logiciel peut être d'une grande aide à cet instant de la recherche. En effet, le logiciel ATLAS.ti permet de réaliser des schémas à partir de codes, citations, mémos, ou concepts développés lors des phases précédentes de l'analyse. La représentation graphique des liens entre les éléments peut être d'une grande aide dans l'identification de la catégorie centrale. Aussi, le logiciel peut fournir une indication car il permet d'avoir accès au nombre de liens établis autour de chacun des codes créés. En outre, l'enregistrement du processus d'interprétation au fur et à mesure que la recherche se développe aide le chercheur à se sensibiliser vers l'identification de la catégorie centrale.

Si ces outils graphiques et logiques ne donnent pas de réponse incontestable à la question de la centralité (les liens ont été établis par le chercheur au cours de son analyse), elle permet néanmoins de guider le chercheur dans l'établissement de la catégorie centrale. La figure 4 reprend l'exemple développé précédemment à un stade postérieur de l'analyse. Dans ce schéma, on a élagué les thèmes et analyses connexes pour se centrer sur la centralité de l'analyse qui semble être le lien entre les caractéristiques du projet et le rôle à jouer par le chef de projet. Ici, seuls les mémos où sont produites les analyses et regroupés certains thèmes

ont été reliées. Les codes s'ont pas été mis sur le schéma afin d'éviter l'encombrement visuel. L'étape ultime est alors la restitution de la recherche selon le format choisi.

Figure 4 : Illustration graphique de recherche de centralité



Weitzman (2000), Lee et Esterhuizen (2000) et Lacey et Luff (2001) synthétisent les bénéfices de l'utilisation de ce type de logiciel :

- vitesse et flexibilité : il permet d'alléger le chercheur du travail « mécanique » de la recherche qualitative (organiser, chercher, récupérer, représenter) et de tester des interprétations différentes (exploration) ;
- représentation : il rend plus tangible le processus de recherche dans la mesure où la visualisation graphique aide à communiquer les résultats ;
- cohérence interne : il facilite l'exploration des données et permet de coordonner le travail d'une équipe de chercheurs.

3. PIEGES LIES A L'UTILISATION DES LOGICIELS POUR L'ANALYSE DES DONNEES QUALITATIVES

L'utilisation des logiciels de support à l'analyse qualitative de données peut elle-même devenir un piège dans la mesure où le chercheur ne se rend pas compte que le logiciel ne fera pas le travail d'analyse. Au-delà de cette incompréhension, existe aussi l'idée reçue que

l'utilisation d'un logiciel dans le cadre d'un projet de recherche apporterait la légitimité nécessaire au chercheur dans la communauté scientifique. Le piège est justement d'avoir l'illusion que le logiciel résoudra le problème d'interprétation et que quelque évaluateur mal informé puisse croire qu'une recherche développée à l'aide d'un logiciel sera forcément de bonne qualité. Ainsi, Cusin (2006) lauréat du prix de thèse Gaëtan Pirou de la Chancellerie des Universités, spécialisé en Sciences Economiques et Gestion en 2007, a utilisé un codage manuel (donc sans logiciel) pour mener sa recherche. La même année, Joffre (2006), lauréat prix de thèse FNEGE-AIMS "Management stratégique", a pour sa part utilisé le logiciel CAQDAS nommé Nvivo.

En plus de cette confusion concernant le rôle du logiciel dans le projet, la démarche d'analyse de la théorie enracinée à l'aide du logiciel cache plusieurs pièges. Les pièges sont des limites latentes et difficiles à identifier aux yeux d'un chercheur inexpérimenté qui peut voir la qualité de sa recherche amoindrie s'il ne peut éviter certains écueils pourtant discernables *a priori*. Nous avons identifié cinq pièges dans le processus d'analyse des données au travers de la théorie enracinée en nous basant sur notre expérience et les travaux de Dembkowski et Hanmer-Lloyd (1995), Weitzman (2000), Lee et Esterhuizen (2000), Lacey et Luff (2001), Gilbert (2002), Kelle (2002) et Bauer et Gaskell (2000).

3.1. Piège 1 : Le chercheur peut s'éloigner de la réalité des sujets

L'influence des logiciels CAQDAS sur le concept de *closeness*, ou la mesure de proximité du chercheur par rapport à la réalité des sujets de la recherche, a été investiguée par Gilbert (2002) par une étude empirique. L'auteur a identifié trois aspects du processus d'éloignement qui surviennent au fur et à mesure que les utilisateurs se familiarisent avec le logiciel. Le premier est la rupture tactile-numérique qui fait référence au changement du travail du chercheur, traditionnellement sur le papier, vers le travail directement sur l'écran d'ordinateur. Celui-ci, limité par sa taille, empêche le chercheur de manipuler librement les différentes pièces de données qu'il a en sa possession. Bournois *et al.* (2002) mentionnent dans cette lignée le risque de présenter un trop grand nombre de codes dans l'analyse et la restitution des données, en particulier dans la lisibilité d'une schématisation graphique.

La gestion du processus de codage et la récupération automatique de l'information demandée est une fonctionnalité du logiciel qui produit les bénéfices de vitesse et flexibilité. Néanmoins, cette apparente facilité cache le deuxième aspect du processus d'éloignement que Gilbert (2002) a appelé le « piège du codage ». Au cours de l'étape du codage ouvert, le chercheur doit produire des codes, microanalyse et mémos qui serviront de base à son

analyse. Avec le logiciel, le nombre de ces éléments tend à augmenter, ce qui par conséquent peut freiner, voire paralyser, le chercheur et l'empêcher de bien réfléchir sur l'essence des données. D'ailleurs, coder se traduit de façon pratique par fracturer les données. Plus on code plus on détache les données de leurs contextes originaux des occurrences. Bien que les logiciels essayent de toujours récupérer les informations dans leurs contextes, une fracturation excessive peut supplanter la capacité humaine de réfléchir sur les concepts cachés derrière les codes. Quelques logiciels, tel qu'ATLAS.ti, ont un outil de liens hypertextes qui gère la construction d'une structure logique à travers les données. Ces liens relient différents segments par des associations définies par le chercheur. Ici encore, conscient de ce piège, le chercheur pourra utiliser au mieux les possibilités du logiciel pour améliorer sa recherche.

Le troisième aspect de l'éloignement, identifié par Gilbert (2002) dans son étude empirique, est le changement métacognitif. Il s'agit de la reproduction de la façon de travailler du logiciel par le chercheur lors de ses activités d'interprétation. Le chercheur n'est alors plus fidèle à la réalité des sujets dont il traite, mais fonctionne selon l'algorithme de recherche du logiciel. Il crée les codes en ayant à l'esprit, non pas la façon au travers de laquelle les sujets de la recherche interprètent le monde, mais comment les codes pourront être réutilisés ultérieurement grâce au logiciel. Pendant, le codage ouvert le chercheur peut se détourner de la réalité des sujets et opter pour la maximisation de l'usage des outils de recherche, et ne revenir aux données qu'ultérieurement, lors de la validation empirique des propositions. Cet aspect est particulièrement dangereux pendant le codage axial, car les outils de recherche, par exemple, l'outil *query tool* d'ATLAS.ti, peuvent faire émerger de façon numérique, de par la puissance de mise en relation possible sur une base de données très large, des liens qui ne reflètent pas les relations empiriques entre codes.

3.2. Piège 2 : Le logiciel peut induire à une analyse superficielle

Ce deuxième piège est lié au premier puisque l'éloignement de la réalité des sujets implique une analyse superficielle. Mais, l'inverse n'est pas forcément vrai. Un chercheur qui ne serait pas éloigné de la réalité des sujets peut produire des résultats superficiels quand :

- 1) il bascule vers un codage factuel ;
- 2) il réifie les codes ;
- 3) il est incité par le logiciel à traiter un échantillon dont la taille supplante ses ressources de temps et énergie.

Le basculement involontaire vers le codage factuel est motivé par la facilité de coder et récupérer l'information. Le chercheur tend à mécaniser le processus de codage au détriment

de la lecture, la contemplation et la réflexion sur les données. Cette pratique s'apparente à ce que Richards (2002a) a appelé le fétichisme du codage. Le codage est trop valorisé au détriment de la découverte. Cela empêche l'émergence de catégories pertinentes, point fort de la technique d'analyse de la théorie enracinée. Les codes créés ne sont que des artefacts du chercheur sans sens pour les sujets (Kelle, 2002).

La réification des codes (Seidel et Kelle, 1995) désigne le phénomène où le chercheur commence à penser que les codes existent réellement et que leur manipulation vers la construction du modèle lui suffit pour avancer dans la recherche, en oubliant les sujets. Malgré les outils graphiques des logiciels CAQDAS comme ATLAS.ti, qui pourraient atténuer le problème de traiter un nombre de codes élevé, cet écueil reste dangereux et peut résulter en une analyse superficielle. Au fur et à mesure que le chercheur forme les réseaux de codes et les représente dans des schémas graphiques séduisants et « complexes », il peut avoir l'impression que cela lui suffise pour que le projet de recherche soit sur le point d'aboutir. C'est l'effet « il est impossible que ce ne soit pas le bon résultat ».

Aussi, le chercheur peut être tenté d'insérer dans son programme de recherche un grand nombre de données, au sein d'une base intégrée et gérée par le logiciel. Au premier abord, il semble que l'extension des données dans le but de la mise en œuvre des comparaisons constantes soit une bonne chose. Néanmoins, une recherche a toujours une durée limitée dans le temps, et s'échelonne selon les impératifs du chercheur et/ou de l'institution au sein de laquelle il réalise sa recherche. Ainsi, le traitement d'un volume trop important de données peut amener le chercheur à vouloir tout intégrer, et de manquer de finesse dans son analyse. Goulding (2001) recommande une analyse des textes ligne par ligne. Si un entretien individuel de soixante minutes avec un acteur comporte plus de sept cent lignes, que des recherches comptent parfois soixante à soixante-dix entretiens, parfois d'une durée qui peut dépasser trois heures, que d'autres documents sont intégrés à la base de données (e-mails, rapports d'activités, courriers, etc.), on comprend aisément que le chercheur n'aura pas le temps matériel d'approfondir son analyse tel que l'exigerait une application stricte des recommandations de la théorie enracinée, et l'amène à proposer une analyse superficielle.

3.3. Piège 3 : Un logiciel n'est pas méthodologiquement neutre

La flexibilité méthodologique des logiciels CAQDAS permet leur utilisation dans le cadre de différentes stratégies de recherche, et le chercheur pourra utiliser les logiciels selon ses besoins spécifiques (Barry, 1998 ; Bournois *et al.*, 2002). Pourtant, le chercheur doit s'adapter aux fonctions, ressources et logiques du logiciel, sachant que les logiciels sont plus ou moins

influencés par l'orientation méthodologique des chercheurs créateurs, comme ATLAS.ti avec son adaptation particulière à la théorie enracinée. Le logiciel n'est donc pas neutre et son usage entraîne des effets sur les décisions méthodologiques du projet de recherche (Weitzman et Miles, 1995). Si le chercheur ne doit pas se laisser guider par le logiciel, il lui faut toutefois lui être fidèle dans ses orientations épistémologiques et méthodologiques (Weitzman, 2000). Pour le cas d'ATLAS.ti, Muhr (1991) avertit que bien que le logiciel soit flexible pour accommoder des stratégies de recherche diverses, ATLAS.ti matérialise une conception méthodologique en soi.

Le chercheur doit alors produire une pensée réflexive sur sa démarche afin de prendre conscience des effets produits par le logiciel sur son travail de recherche. De telles analyses sont trop rares à l'heure actuelle dans la démarche de restitution des résultats de la recherche (Langley et Royer, 2006). Si la démarche de Malina et Selto (2001), où l'ensemble des étapes de l'utilisation du logiciel est décrite, est une première étape louable et nécessaire, les auteurs laissent toutefois l'impression que le logiciel est un organisateur de données méthodologiquement neutre.

3.4. Piège 4 : Le rapport coût / bénéfice est difficile à évaluer

Avant le travail d'analyse, de nombreux chercheurs se posent différentes questions quant à l'utilisation d'un logiciel CAQDAS. Cela vaut-il la peine d'investir dans l'apprentissage du logiciel ? Cette connaissance sera-t-elle utile pour des projets futurs ? Quel logiciel l'équipe de chercheurs va-t-elle retenir ? La décision d'utiliser le logiciel dans le projet de recherche doit considérer les coûts d'apprentissage, les coûts de préparation de la base de données, ainsi que le temps disponible dans l'horizon du projet. Il est difficile d'opter pour l'utilisation d'un logiciel quand le chercheur manque de temps dans le planning du projet, ou si le logiciel risque de n'être utilisé qu'une seule fois.

Pourtant, il y a quelques stratégies pour faire diminuer les coûts d'applications des logiciels. Le chercheur peut décider de n'utiliser que quelques fonctions choisies du logiciel, ce qui réduit les coûts d'apprentissage et d'engagement. Par exemple il peut utiliser le logiciel uniquement pour représenter graphiquement les résultats qui ont été obtenus par la façon traditionnelle « papier-crayon ». Une solution possible pour accélérer l'apprentissage du logiciel réside dans les rencontres en groupe. Il s'agit de séances d'analyse, coordonnées par un chercheur plus expérimenté dans l'usage du logiciel, à laquelle participent les chercheurs en cours d'analyse à l'aide du logiciel. Chaque personne présente ses interprétations et les derniers progrès réalisés dans l'analyse et le maniement du logiciel pour que les autres

puissent critiquer, remettre en cause, ou suggérer des pistes dans l'usage du logiciel. Cela permet aussi d'améliorer la cohérence interne des résultats par échange d'interprétations.

3.5. Piège 5 : Un logiciel CAQDAS transforme un chercheur incompetent en chercheur encore plus incompetent

Ce dernier piège est peut-être le plus néfaste. La croyance dans la fausse capacité du logiciel à interpréter et analyser les données peut précipiter certains chercheurs inexpérimentés dans la recherche qualitative. Pis encore, cette croyance peut décourager les jeunes chercheurs à suivre une bonne formation d'un point de vue méthodologique. Par conséquent, un chercheur incompetent dans la stratégie de recherche choisie verra ses déficiences augmenter par l'utilisation du logiciel parce qu'il ne sera pas capable d'éviter les pièges ici décrits. Le « côté humain » de la recherche qualitative à l'aide d'un logiciel est fondamental depuis la collecte de données. Le logiciel n'a rien à offrir à un chercheur incompetent qui analyse un entretien ou une base de données pauvre. Le logiciel ne doit ainsi pas laisser penser qu'il se subordonne aux réflexions analytiques et méthodologiques nécessaires à toute recherche de qualité.

4. CONSIDERATIONS FINALES

Les principales critiques faites aux recherches basées sur une analyse qualitative des données vont dans le sens d'un manque de rigueur quand à l'analyse et à la présentation des résultats des recherches. La théorie enracinée et l'utilisation de logiciel sont souvent citées comme des stratégies permettant d'améliorer la qualité et la rigueur de la recherche. Les deux sont fréquemment associées l'une avec l'autre et partagent autant de bénéfices que d'inconvénients. Nous avons alors expliqué la conduite d'une recherche basée sur la théorie enracinée en combinant les règles de cette méthode avec l'utilisation d'un logiciel CAQDAS, qui permet de développer de nombreux éléments de cette méthode en référençant et en organisant l'ensemble des éléments nécessaires à un construit théorie de qualité : catégories, mémos, codes, réseaux de thèmes, etc.

Néanmoins, l'utilisation des logiciels d'analyse des données qualitatives recouvre deux volets : l'un représentant les promesses des fonctionnalités d'un tel outil, l'autre érigeant des écueils qui attendent le chercheur tout au long de sa recherche. Nous avons montré que la prise de conscience de ces deux versants de l'utilisation des logiciels CAQDAS dans l'optique de la théorie enracinée est une des conditions nécessaires pour que les logiciels apportent

qualité et rigueur au sein de la théorie enracinée : profiter de avantages potentiels mais aussi éviter les pièges.

D'après Lee et Fielding (1991), les logiciels souffrent du « syndrome de Frankenstein » : malgré son apparence, et considéré comme un monstre par les autres, il a de bonnes intentions et le cœur sur la main. L'incompréhension et la mauvaise utilisation du logiciel sont ainsi les causes principales de la formation d'une image négative à propos des logiciels CAQDAS. Ces travers peuvent être rencontrés majoritairement chez de jeunes chercheurs qui n'ont pas eu l'habitude de développer leur réflexion sans ces logiciels, comme chez des chercheurs plus aguerris qui ne domineraient pas la technologie et se laisseraient guider par des préjugés. Ces pièges ont été mis en évidence dans l'optique de la théorie enracinée mais peuvent se retrouver pour d'autres stratégies. En discutant des possibilités et pièges, nous espérons que les chercheurs se forment une opinion réfléchie à propos de tels logiciels et de leur usage, et prennent la bonne décision concernant l'usage des logiciels en fonction de leur projet de recherche et de leur sensibilité à l'outil.

REFERENCES

- ABEL T., (1948), «The Operation Called Verstehen», *The American Journal of Sociology*, Vol. 54, No. 3, pp.211-218.
- ALEXA M. et ZUELL C., (1999), *Commonalities, differences and limitations of text analysis software: the results of a review*, http://193.175.239.100/Publikationen/Berichte/ZUMA_Arbeitsberichte/99/99_06.pdf.
- BARRY C. A., (1998), «Choosing Qualitative Data Analysis Software: Atlas/ti and Nudist Compared», *Sociological Research Online*, Vol. 3, No. 3.
- BAUER M. et GASKELL G. (eds), (2000), *Qualitative researching with text, image and sound*, Sage.
- BERGER P. et LUCKMANN T., (1966), *The social construction of reality*, Penguin Books.
- BOURNOIS F., POINT S. et VOYNET-FOURBOUL C., (2002), «L'analyse des données qualitatives assistée par ordinateur : une évaluation», *Revue Française de Gestion*, Vol. 137.
- CHARMAZ K., (2000) «Grounded Theory: Objectivist and Constructivist Methods», dans DENZIN N. K. et LINCOLN Y. S. (eds), *Handbook of Qualitative Research*, Thousand Oaks, pp.509-535.
- CHARMAZ K., (2006), *Constructing grounded theory: a practical guide through qualitative analysis*, Sage.
- CICOURCEL A., (2003) «Contre un empirisme naïf. Une théorie plus forte et un contrôle plus ferme sur les données», dans CEFAÏ D. (ed.) *L'enquête de terrain*, La Découverte.
- CORBIN J. et STRAUSS A., (1990), «Grounded Theory Research: Procedures, Canons, and Evaluative Criteria», *Qualitative Sociology*, Vol. 13, No. 1, pp.3-21.
- CUSIN J., (2006), *L'apprentissage par l'échec commercial*, thèse de doctorat, CREPA, Université Paris Dauphine.
- DAVID A., (1999), «Logique, épistémologie et méthodologie en sciences de gestion», *communication à Association Internationale de Management Stratégique*, Chatenay-Malabry.

DEMBROWSKI S. et HANMER-LLOYD S., (1995), «Computer applications - a new road to qualitative data analysis?», *European Journal of Marketing*, Vol. 29, No. 11, pp.50-62.

DOUGLAS D., (2003), «Grounded Theories of Management: A methodological Review», *Management Research News*, Vol. 26, No. 5, pp.44-52.

FRIEDBERG E., (1994) «Le raisonnement stratégique comme méthode d'analyse et comme outil d'intervention», dans PAVÉ F. (ed.) *L'analyse stratégique. Sa genèse, ses applications et ses problèmes actuels. Autour de Michel Crozier*, Éditions du Seuil, pp.135-152.

GIBBS G. R., (2002), *Qualitative Data Analysis. Explorations with Nvivo*, Open University Press, *Understanding Social Research*.

GILBERT L., (2002), «Going the distance: 'closeness' in qualitative data analysis software», *International Journal of Social research Methodology*, Vol. 5, No. 3, pp.215-228.

GLASER B. G., (1978), *Theoretical Sensitivity*, The Sociology Press.

GLASER B. G., (1992), *Basics of Grounded Theory Analysis: Emergence vs. Forcing*, The Sociology Press.

GLASER B. G., (2002), «Constructivist Grounded Theory?», *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, Vol. 3, No. 3.

GLASER B. G. et STRAUSS A., (1967), *The discovery of grounded theory*, Adline.

GOULDING C., (2001), «Grounded Theory: A Magical Formula or a Potential Nightmare», *The Marketing Review*, Vol. 2, No. 1, pp.21-34.

GOULDING C., (2002), *Grounded Theory: A Practical Guide for Management, Business and Market Researchers*, Sage.

HUBER G. et GARCIA C., (1991), « Computer assistance for testing hypotheses about qualitative data: the software package AQUAD 3.0», *Qualitative Sociology*, Vol. 14, No. 4, pp.325-347.

JOFFRE O., (2006), *Management de l'intégration dans les « fusions entre égaux »*, thèse de doctorat, CREPA, Université Paris Dauphine.

KAUFMANN J.-C., (2007), *L'entretien compréhensif*, Armand Colin.

KELLE U., (2002) «Análise com auxílio de computador: codificação e indexação», dans BAUER M. et GASKELL G. (eds), *Pesquisa qualitativa com texto, imagem e som: um manual prático*, Vozes.

KELLE U., (2005), «"Emergence" vs. "Forcing" of Empirical Data? A Crucial Problem of "Grounded Theory" Reconsidered», *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, Vol. 6, No. 2.

LACEY A. et LUFF D., (2001), *Qualitative data analysis*, Trent Focus.

LANGLEY A. et ROYER I., (2006), «Perspectives on Doing Case Study Research in Organizations», *M@n@gement*, Vol. 9, No. 3, pp.73-86.

LEE R. et FIELDING N., (1991) «Computing for qualitative research: options, problems and potential», dans FIELDING N. et LEE R. (eds), *Using computers in qualitative research*, Sage.

LEE R. M. et FIELDING N., (1996), «Qualitative Data Analysis: Representations of a Technology: A Comment on Coffey, Holbrook and Atkinson», *Sociological Research Online*, Vol. 1, No. 4.

LEE R. M. et ESTERHUIZEN L., (2000), «Computer software and qualitative analysis: trends, issues and resources», *International Journal of Social Research Methodology*, Vol. 3, No. 3, pp.231-243.

LOCKE K., (2001), *Grounded theory in management research*, Sage.

LONKILA M., (1995) «Grounded Theory as an Emerging Paradigm for Computer-Assisted Qualitative Data Analysis», dans KELLE U. (ed.) *Computer-Aided Qualitative Data Analysis: Theory, Methods and Practice*, Sage, pp.41-51.

- MALINA M. A. et SELTO F. H., (2001), «Communicating and Controlling Strategy: An Empirical Study of the Effectiveness of the Balanced Scorecard», *Journal of Management Accounting Research*, Vol. 13, No. 1, pp.47-90.
- MERRIAM S. B., (1998), *Qualitative research and case study applications in education*, Jossey-Bass.
- MILES M., (1979), «Qualitative data as an attractive nuisance. The problem of analysis», *Administrative Science Quarterly*, Vol. 24, No. 4, pp.590-601.
- MORAN D., (2002), *Introduction to phenomenology*, Routledge, 2ème ed.
- MORGAN G. et SMIRCICH L., (1980), «The case for qualitative research», *Academy of Management Review*, Vol. 5, No. 4, pp.491-500.
- MUHR T., (1991), «ATLAS/ti: a prototype for the support of text interpretation», *Qualitative Sociology*, Vol. 14, No. 4.
- RICHARDS L., (2002a), «Qualitative Computing - A methods revolution?», *International Journal of Social Research Methodology*, Vol. 5, No. 3, pp.263-276.
- RICHARDS T., (2002b), «An intellectual history of NUD*IST and Nvivo», *International Journal of Social Research Methodology*, Vol. 5, No. 3, pp.199-214.
- RICHARDS T. et RICHARDS L., (1991), «The NUDIST qualitative data analysis system», *Qualitative Sociology*, Vol. 14, No. 4, pp.307-324.
- SEIDEL J. et KELLE U., (1995) «Different functions of coding in the analysis of textual data», dans KELLE U. (ed.) *Computer-aided qualitative data analysis: theory, methods and practice*, Sage.
- SIEDEL J. et CLARK J., (1984), «THE ETHNOGRAPH: A computer program for the analysis of qualitative data», *Qualitative Sociology*, Vol. 7, No. 2, pp.110-125.
- STRAUSS A. et CORBIN J., (1990), *Basics Qualitative Research: Grounded Theory Procedures and Techniques*, Sage.
- STRAUSS A. et CORBIN J., (1998), *Basics of Qualitative Research. Techniques and Procedures for Developing Grounded Theory*, Sage, 2ème ed.
- STRÜBNG J., (1997) «Computer Tools for Grounded Theory: Introducing ATLAS/ti for Windows 95», dans KLAR R. et OPTZ O. (eds), *Classification and Knowledge Organization: Proceedings of the 20th Annual Conference of the Gesellschaft Fur Klassifikation*, Springer.
- TURNER B. A., (1983), «The Use of Grounded Theory for the Qualitative Analysis of Organizational Behaviour», *Journal of Management Studies*, Vol. 20, No. 3, pp.333-348.
- WEITZMAN E., (2000) «Software and qualitative research», dans DENZIN N. K. et LINCOLN Y. S. (eds), *Handbook of qualitative research*, Sage.
- WEITZMAN E. et MILES M., (1995), *Computer programs for qualitative data analysis: a software sourcebook*, Sage.
- WELLS K., (1995), «The strategy of grounded theory: Possibilities and problems», *Social Work Research*, Vol. 19, No. 1, pp.33-37.