

Cours 12 - Éléments empiriques (4/4) (3 avril 2024)

Alignement pédagogique

Mise en application

Objectifs visés et activités associées

Objectif général : 4. Appliquer les méthodes de base en sciences de l'information pour analyser des données

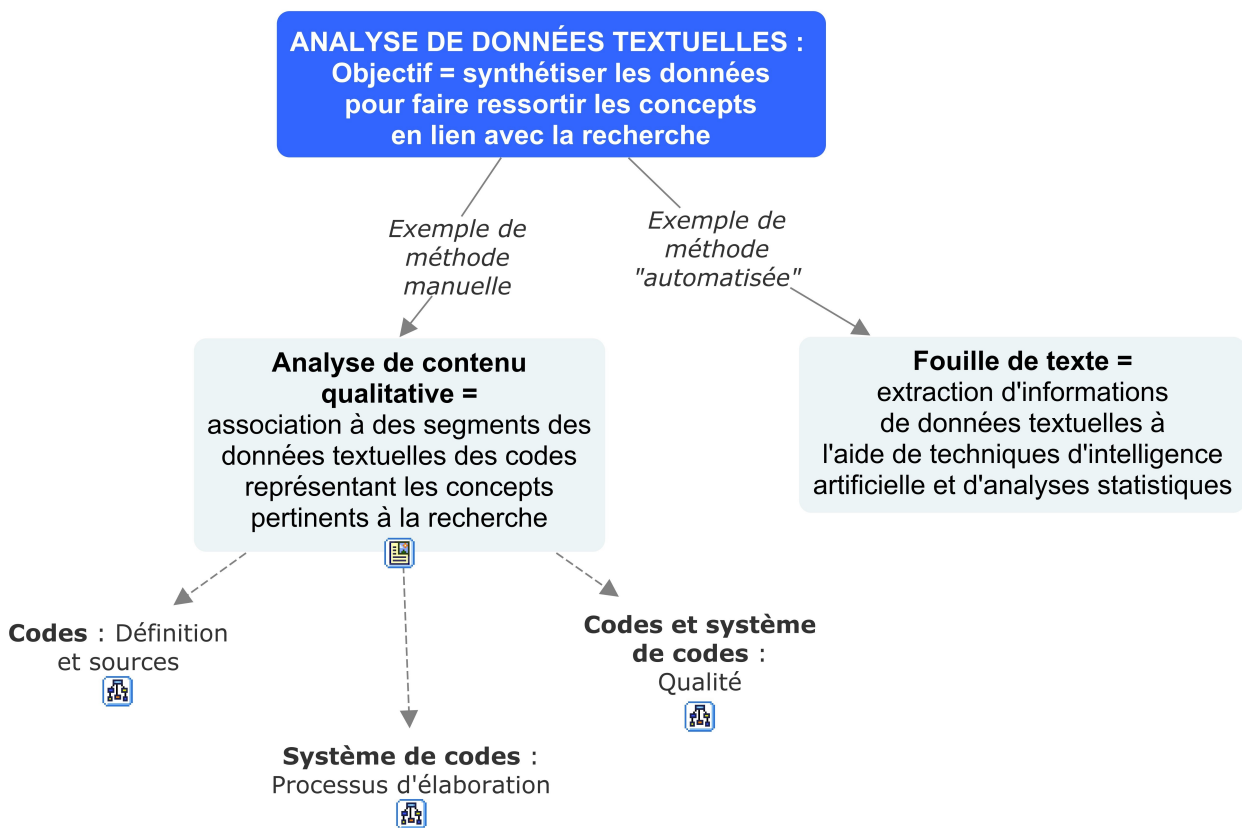
Objectif spécifique : 4a. Décrire les méthodes et outils d'analyse pour des données qualitatives

Activité : TP3

Cours 12 - Éléments empiriques (4/4) (3 avril 2024)

Cartes conceptuelles

Phase 3 : Éléments empiriques : Analyse de données textuelles



© Christine Dufour 2005-2018, 2024, EBSI, Université de Montréal
(Modifications Virginie Wenglenski et Marielle St-Germain © 2023)

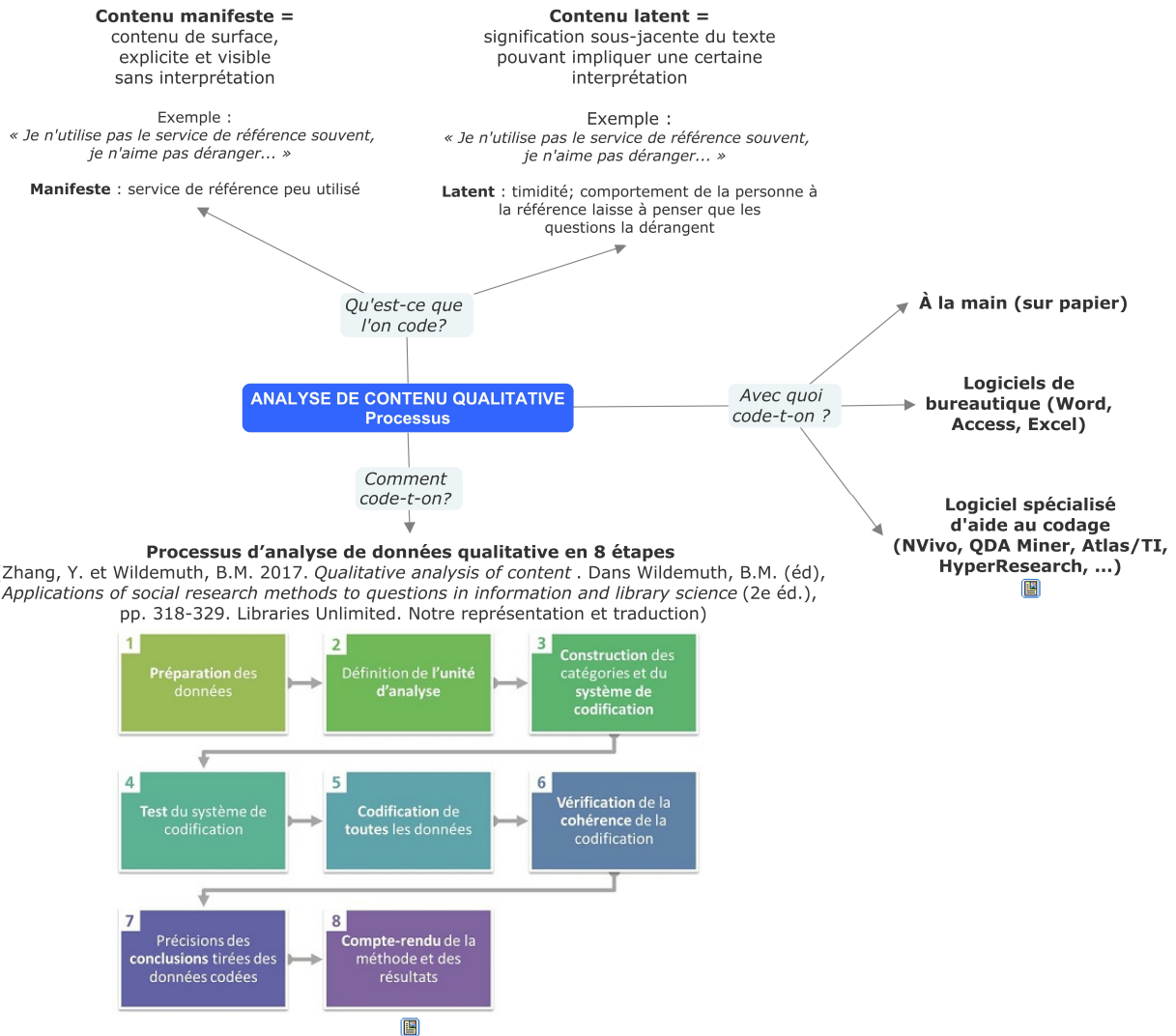
https://reseauconceptuel.umontreal.ca/rid=1YRFSB0P5-2JWWL6-9KST/sci6007_c13_pe_analyse_contenu.cmap

Synopsis

Parmi les données recueillies par certains types de recherche on retrouve parfois des données textuelles plus ou moins structurées. C'est le cas, par exemple, pour une étude de cas se basant principalement sur des entrevues semi-structurées ou sur des questions ouvertes dans un questionnaire pour une enquête. Différentes méthodes existent afin de pouvoir synthétiser ces données dont des méthodes manuelles comme

l'analyse de contenu qualitative ou la fouille de texte pour une approche plus quantitative et plus automatisée utile notamment pour de grands corpus de texte.

Phase 3 : Éléments empiriques > Analyse de contenu qualitative > Élaboration du système de codes



© Christine Dufour 2005-2018, 2024, EBSI, Université de Montréal
(Modifications Virginie Wenglenski et Marielle St-Germain © 2023)

https://reseauconceptuel.umontreal.ca/rid=1YRFSB0P5-2LB33R-9KRC/sci6007_c13_pe_analyse_contenu_contenu.cmap

Synopsis

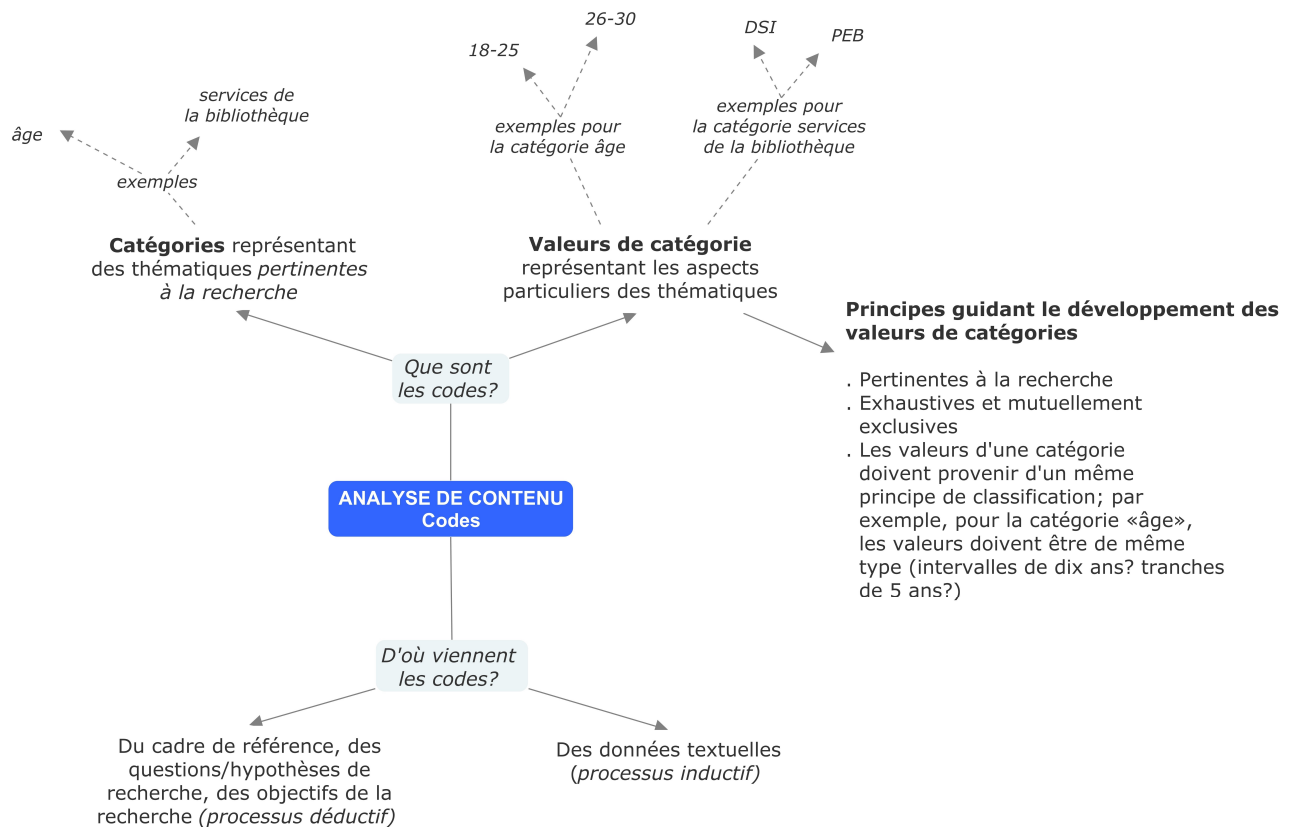
L'analyse de contenu qualitative peut s'intéresser à deux types de contenu. D'une part, il y a le contenu manifeste, c'est-à-dire ce qui est visible sans aucune interprétation. D'autre part, on peut aussi s'intéresser au contenu latent qui s'attardera à la signification sous-jacente du contenu ce qui peut impliquer une certaine interprétation.

Il n'existe pas une seule manière de faire de l'analyse de contenu qualitative. Plusieurs processus existent se rattachant entre autres à différentes approches et disciplines. Zhang et Wildemuth proposent un processus en 8 étapes allant de la préparation des données (par exemple en retranscrivant des enregistrements audio), à la définition de l'unité d'analyse textuelle (le mot? la phrase? le paragraphe?), à la construction d'un

premier arbre de codage qui, par la suite est testé. Une fois le test concluant, ce dernier est appliqué à toutes les données pour, au besoin, faire une autre vérification de la cohérence de la codification effectuée. Les deux dernières étapes consistent à tirer des conclusions du codage effectué et à consigner et diffuser la méthode et les résultats.

Le codage pourrait se faire à la main sur des impressions des données textuelles. Des logiciels bureautiques pourraient aussi être utilisés pour cette tâche, que ce soit Word ou Excel par exemple. Finalement, des logiciels dédiés à ces tâches existent. Ces derniers ne feront pas le travail de codage comme tel, mais offriront des fonctionnalités pour aider le chercheur ou la chercheuse à le faire et à exploiter le codage effectué.

Phase 3 : Éléments empiriques > Analyse de contenu qualitative > Codes



© Christine Dufour 2005-2018, 2024, EBSI, Université de Montréal (Modifications Virginie Wenglenski et Marielle St-Germain © 2023)

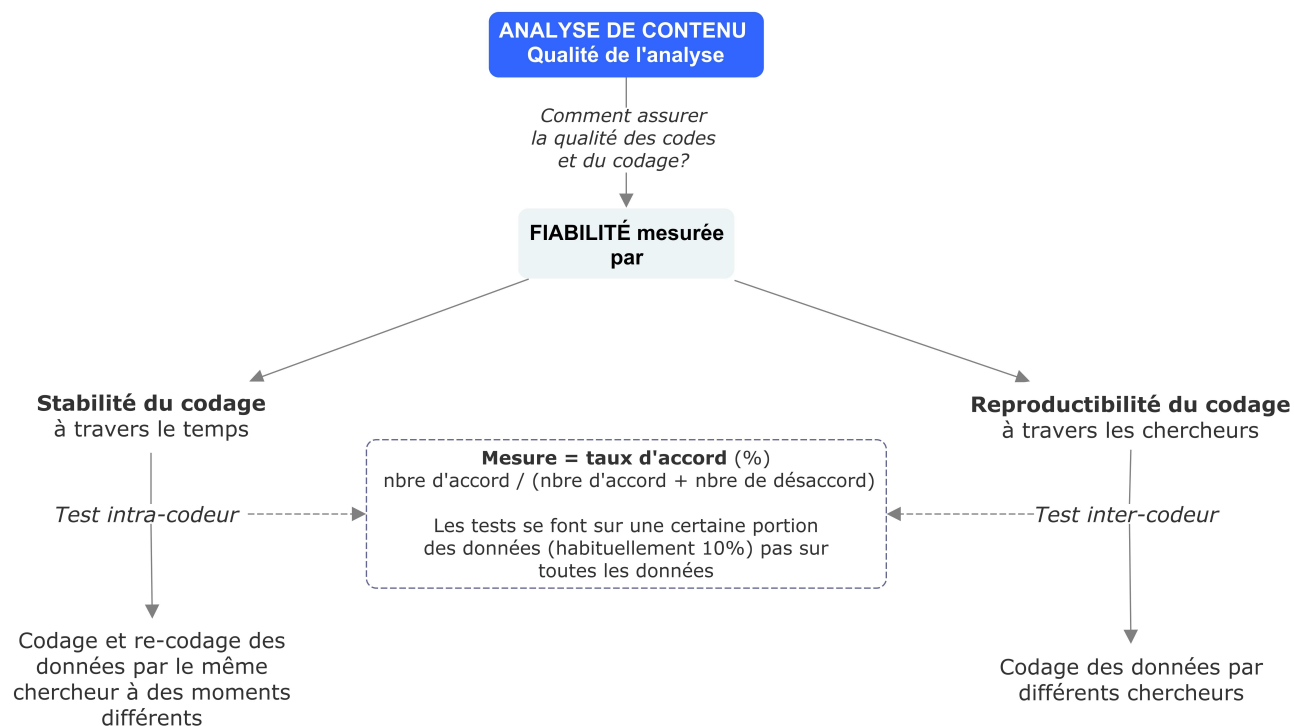
https://reseauconceptuel.umontreal.ca/rid=1YRFSB0P5-P9WM1J-9KQZ/sci6007_c13_pe_analyse_contenu_codes.cmap

Synopsis

Certaines approches préconisent deux niveaux dans le codage. Au premier niveau, on définit des catégories qui représentent les thématiques couvertes dans les données. Par exemple, on pourrait avoir une catégorie "services de la bibliothèque". Ces catégories se déclinent par la suite selon leurs différentes valeurs que l'on retrouve dans les données. La catégorie "services de la bibliothèque" pourrait, par exemple, avoir comme valeurs "service de référence", "service de prêt", etc. Ces valeurs doivent notamment être pertinentes à la recherche ainsi que mutuellement exclusives.

Dans une approche qualitative, les codes viennent principalement des données. On parle en ce cas de codes inductifs qui émergent des données. Il est aussi possible d'avoir certains codes qui proviennent par exemple

du cadre de référence de la recherche, soit des codes déductifs.



© Christine Dufour 2005-2018, 2024, EBSI, Université de Montréal
(Modifications Virginie Wenglenski et Marielle St-Germain © 2023)

https://reseauconceptuel.umontreal.ca/rid=1YRFSB0P5-3XWJ03-9KSC/sci6007_c13_pe_analyse_contenu_qualite.cmap

Synopsis

Afin d'augmenter la qualité de la recherche il importe de prendre les moyens nécessaires pour assurer la qualité de l'analyse de contenu qualitative effectuée. La fiabilité du codage peut se mesurer entre autres par un test intra-codeur où le chercheur ou la chercheuse fait le codage une première fois et, après un laps de temps assez long pour avoir oublié ce qui a été fait, refait le codage une deuxième fois. Les deux analyses sont alors comparées pour vérifier la cohérence du codage. On vise souvent au moins 90% d'accord (un accord étant un même code appliqué à un même extrait). Si ce taux n'est pas atteint, il faut comprendre les raisons derrière les désaccords pour apporter des correctifs. Un nouveau test intra-codeur aura ensuite lieu pour vérifier que les problèmes ont bien été corrigés.

La fiabilité du codage peut aussi se mesurer par un test inter-codeur où, cette fois, deux personnes différentes codent des données, codage qui est ensuite comparé. On recherchera par exemple un taux d'accord d'au moins 80%. S'il n'est pas atteint, les deux personnes ayant fait le codage discutent pour identifier les sources de divergence, des modifications sont apportées et un nouveau test inter-codeur est effectué. Il est à noter que, dans certains contextes, il peut être justifié de ne pas faire de test inter-codeur. C'est le cas, par exemple, d'une recherche où le chercheur ou la chercheuse principale est en immersion prolongée dans son terrain et que le contenu latent est codé. Il est difficile, voire impossible pour une autre personne qui n'est pas en immersion dans le terrain d'arriver aux mêmes interprétations.