

## Introduction (critique) de QDA miner, logiciel d'Analyse Qualitative des données. Application à la traduction en Français de sourate Al Baqara par Hamidullah

Ali Kouaouci

Université de Montréal- Canada

[ali.kouaouci@umontreal.ca](mailto:ali.kouaouci@umontreal.ca)

### Résumé de l'article

L'objectif de cet article est double : présenter concrètement le logiciel d'analyse qualitative QDA miner et tester sa capacité à identifier les principaux thèmes de la Sourate Al Baqara, le plus long chapitre du Saint Coran avec 286 versets traduits en français par le professeur Hamidullah. Les principales caractéristiques de QDA miner seront présentées de façon critique. Plusieurs anomalies ont été identifiées et documentées, dont la plus importante est la faible capacité du logiciel à identifier des thèmes importants.

Même si l'intelligence humaine a été plus convaincante pour identifier des thèmes significatifs, la plupart des auteurs ont opté pour diviser linéairement la sourate par groupement de versets et ce choix a montré ses limites puisque cette façon de découper le texte a laissé échapper deux thèmes primordiaux Allah et Muhammad ﷺ qui sont présents dans toute la Sourate et non concentrés dans un nombre limité de versets.

**Mots-clés :** Analyse qualitative, cooccurrences, similitude, nuage de mots, dendrogramme.

**Valeurs nominales JEL :** C38, C65, C35, C19, Y10

### Abstract:

The objective of this article is twofold: to present in a practical way

the qualitative analysis software QDA miner and to show what can be achieved with this tool to identify the main themes of Sura Al Baqara, the longest chapter of the Holy Quran with 286 verses translated into French by Professor Hamidullah. The main features of QDA miner will be presented critically. Several anomalies have been identified and documented, the most important of which is the software's low ability to identify significant themes.

Even if human intelligence was more convincing in terms of finding significant themes, most authors have opted for dividing the sourate by set of verses and this choice has shown its limitations since this way of cutting out the text has let slip two primary themes Allah and Muhammad ﷺ that are present throughout the Sura and not concentrated in a limited number of verses.

**Keywords:** Qualitative analysis; cooccurrences; similarity; word cloud; dendrogram.

**JEL Classification Codes:** C38, C65, C35, C19, Y10

## Introduction

L'objectif de cet article est double : il s'agit de présenter de manière pratique le logiciel d'analyse qualitative QDA miner développé par Provalis au Québec et de montrer ce qu'on peut réaliser avec cet outil sur un corpus sacré à savoir sourate Al Baqara, le chapitre le plus long du Saint-Coran avec 286 versets traduit en Français par le professeur Hamidullah. On tentera de présenter les principales fonctionnalités de QDA miner de manière critique en identifiant les limites éventuelles. Il y a des spécialistes de l'analyse qualitative qui refusent de recourir à l'analyse quantitative et pourtant lorsqu'un concept se répète 10 fois dans un discours, il y a un message différent par rapport à un concept utilisé une seule fois. Et cela est valable aussi bien dans un sondage sur les conditions de travail que dans un focus-group avec des juges

ou toute autre réponse à une question ouverte. L'idée ici est de démontrer la facilité d'utilisation de ce logiciel et le potentiel qu'il représente pour les chercheurs, tout en montrant certaines de ses limites bien entendu.

Le deuxième objectif est de comparer les principaux thèmes de Al Baqara tels qu'identifiés par QDA miner, autrement dit par l'intelligence artificielle, et tels que formulés par des chercheurs humains. De nombreux spécialistes se sont appliqués à analyser la traduction du Saint-Coran dans différentes langues mais bien entendu l'idéal serait de travailler dans la langue originale l'Arabe. Provalis développe précisément une version arabe de son logiciel. L'idéal serait que des jeunes chercheurs et étudiants puissent maîtriser le logiciel et oser regarder nos textes sacrés avec des outils modernes qui facilitent le travail de recherche pour une meilleure compréhension du message. Notre but ultime est une meilleure compréhension de sourate Al Baqara

De manière générale, l'analyse qualitative des données se doit de respecter le corpus du discours analysé pour restituer autant que possible l'intention de la source de ce discours. Il ne s'agit donc aucunement d'une approche critique du texte mais au contraire d'un effort pour découvrir sa structure et ses principaux thèmes. C'est la démarche des juristes musulmans qui ne cherchent pas à promulguer des nouvelles lois, mais à découvrir la Loi D'Allah en étudiant les textes disponibles (Saint-Coran et Hadiths). Les différents chercheurs qui ont travaillé sur une analyse structurale quantitative et qualitative de Al Baqara sont souvent anglophones, notamment asiatiques. La baqara a été subdivisée en groupes de versets pour identifier les thèmes(OUDA, 2015),(RAHMAN, 2018). Nous aurons l'occasion de comparer les résultats de cette approche avec QDA miner(BARRY R. & MAAS, 2007). La démarche proposée ici tente de dégager une structure en utilisant les aspects quantitatifs et graphiques au lieu de diviser les versets en fonction de leurs thèmes, ce qui est la méthode la plus courante pour tenter de découvrir les thèmes importants abordés dans cette sourate. La fréquence des mots utilisés est en soi indicatrice

des thèmes comme on va le voir. Bien entendu cela ne diminue en rien les travaux qui choisissent d'autres approches : il faudra tirer profit des différentes initiatives pour présenter Al Baqara

Notre approche est en fait basée essentiellement sur les fréquences des concepts qui reviennent le plus souvent, isolément ou accompagné d'autres concepts, désignés justement par QDA miner comme co-occurrences. Dans un deuxième temps, nous montrerons que le codage des versets permettra d'aller au delà de la simple fréquence des mots en démontrant que sans être cité nommément le Prophète Muhammad ﷺ est le Prophète le plus présent dans la sourate.

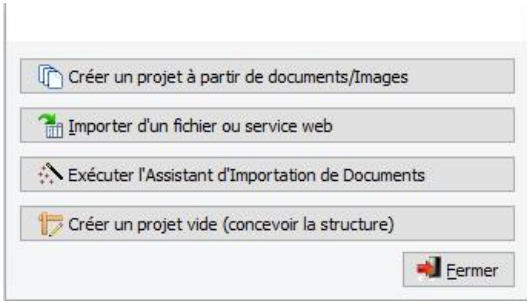
## 1. Méthode

**QDA miner** DA Miner est un logiciel conçu pour la recherche qualitative tout en incorporant la dimension quantitative et avec un grand choix de produits graphiques qui aident à la compréhension des données. La version complète contient WordStat (analyse de contenu) et SimStat (analyse statistique). Il existe une version gratuite QDA Lite qui peut être téléchargée

De nombreuses introductions au logiciel sont disponibles. (COLIN, SIMON, 2010), (<https://qda-miner-lite.software.informer.com/4.0/>, n.d.), (VOYNNET-FOURBOUL, 2012), (QDA miner, Logiciel d'analyse qualitative des données, Manuel de l'utilisateur, 2018)

Attention, la sauvegarde n'est pas automatique et la fonction annuler n'existe pas. En choisissant exécuter l'assistant d'importation, le plus simple est de travailler sur un fichier sous Word de manière à pouvoir isoler les segments. Documents en format RTF, mais la plupart des formats sont acceptés DOC, PDF, Excel, Access, Paradox, dBase, etc.

A l'ouverture de QDA miner, le menu suivant s'affiche



Dans notre cas, la sourate baqara contient 286 versets. Chaque verset doit être précédé de /r pour marquer la séparation. Nous avons gardé les numéros de versets comme numéros de cas.

/r 1. Alif, Lam, Mim.

/r 2. C'est le Livre au sujet duquel il n'y a aucun doute, c'est un guide pour les pieux.

/r 3. Qui croient à l'invisible et accomplissent la Salat et dépensent [dans l'obéissance à Allah], de ce que Nous leur avons attribué.

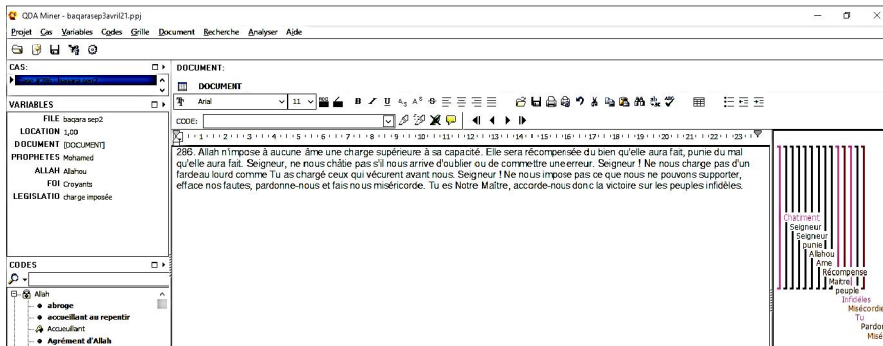
Nous avons ensuite transformé le fichier au format pdf et exécuté l'importation.

En ouvrant notre fichier, on trouve le menu suivant avec chaque verset se trouve dans un tableau pour pouvoir travailler :

- Sur les cas : le verset 3 correspond au cas 3. On peut l'éditer le supprimer, ajouter un cas, etc.
- Sur les variables : si on désire créer des variables comme les noms d'Allah, les Prophètes, on doit pouvoir indiquer pour chaque verset si un des noms d'Allah est cité, ou un de ses Prophètes
- Sur les codes : l'essentiel du travail de l'analyse qualitative réside dans le codage du texte puisque c'est le chemin utilisé pour aboutir à l'analyse thématique. Chaque verset pourra recevoir un code, plusieurs codes ou pas de code du tout, comme par exemple le premier verset Alif.Lam.Mim.

On peut facilement ajouter un code, le supprimer, le modifier ou le fusionner avec d'autres codes, ou le diviser en deux nouveaux codes en cliquant à droite et en choisissant la tâche appropriée.

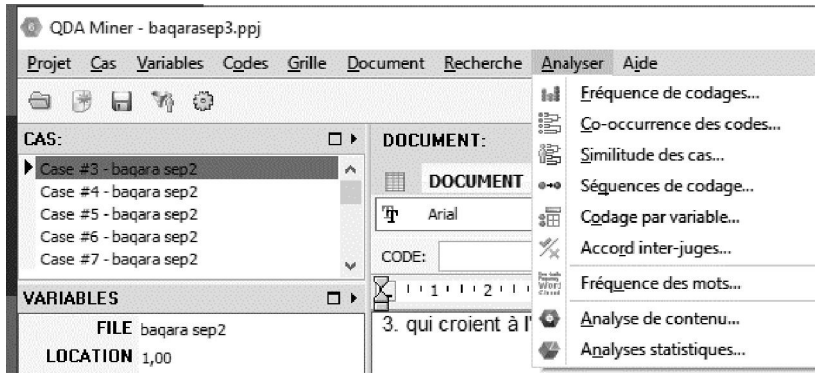
**Fig.1. Page de travail pour coder, ajouter ou supprimer un cas, renseigner la ou les variables titre de la figure**



Dans l'exemple du verset 286, les codes proposés renvoient aux mots clés Seigneur, âme, récompense, pardon, etc. et se trouvent à droite du tableau. Bien entendu, chaque chercheur peut en fonction de son objectif définir d'autres codes.

Une fois le codage achevé, en cliquant sur analyse les choix suivants sont affichés

- Fréquence des codages
- Co-occurrence des codes
- Similitude des cas
- Séquence de codage
- Codage par variable
- Accord inter-juges
- Fréquence des mots
- Analyse de contenu
- Analyse statistique



On notera qu'Analyse de contenu renvoie à WordStat un logiciel associé à QDA miner mais qui peut également être utilisé séparément. Analyse statistique renvoie à SimStat un autre logiciel associé à QDA miner pouvant être utilisé séparément et qui ne sera pas traité dans le cadre de cet article.

L'accord inter-juges ne sera pas abordé ici car cela nécessite différents codeurs travaillant sur le même projet, ce qui n'est pas le cas ici.

Une cooccurrence est la présence simultanée de deux ou de plusieurs codes dans un même énoncé (généralement le paragraphe ou ensemble de paragraphes à déterminer. Les cooccurrences dénotent un lien. Pour calculer la co-occurrence de codes :

- Dans la barre de menu, il faut cliquer sur « Analyser » et sélectionnez « Cooccurrence des codes »
- Sélectionnez les codes et paramétrez le calcul de la cooccurrence avec les zones « Groupement » et « Options d'échelonnement multidimensionnel ».

## • 2, Résultats

### 2.1 Fréquence des codage

Chaque code doit être classé dans une catégorie qui regroupe plusieurs codes. C'est ainsi que nous avons identifié 6 catégories : Allah, la création, les Croyants, les Envoyés, les Incroyants et le

paradis. Sous chacune de ces catégories, les codes ont été choisis pour résumer le mieux possible la signification des versets.

La catégorie Allah a regroupé 23 codes, y compris des pronoms comme Il-je-Moi-Nous-Tes-Toi-Tu qui tous désignent Allah. Bien entendu, ces choix peuvent être remis en question ou améliorés par d'autres chercheurs. Le but ici est purement méthodologique

- Allah
- Création
- Croyants
- Envoyés
- Incroyants
- paradis

**Tableau 1.** Fréquences des codes sous la catégorie Allah

Catégor	Code	Fréquen	%	Ca	% Cas
\	Allah	441	25,60	20	69,90
Allah	Accueillant au	2	0,10%	2	0,70%
Allah	Agrément d'Allah	1	0,10%	1	0,30%
Allah	Défenseur	2	0,10%	1	0,30%
Allah	Divinité	7	0,40%	3	1,00%
Allah	Grâce d'Allah	10	0,60%	10	3,50%
Allah	Il	30	1,70%	22	7,70%
Allah	Indulgent	1	0,10%	1	0,30%
Allah	Je	4	0,20%	4	1,40%
Allah	Le Vivant	2	0,10%	1	0,30%
Allah	Maitre	1	0,10%	1	0,30%
Allah	Moi	11	0,60%	6	2,10%
Allah	Nous	17	1,00%	14	4,90%
Allah	Ordres d'Allah	2	0,10%	1	0,30%
Allah	Prescriptions	2	0,10%	2	0,70%
Allah	Protecteur	2	0,10%	2	0,70%
Allah	Rendre compte	5	0,30%	5	1,70%
Allah	Secoureur	5	0,30%	5	1,70%
Allah	Seigneur	1	0,10%	1	0,30%
Allah	Tes	1	0,10%	1	0,30%
Allah	Toi	8	0,50%	4	1,40%
Allah	Tu	1	0,10%	1	0,30%



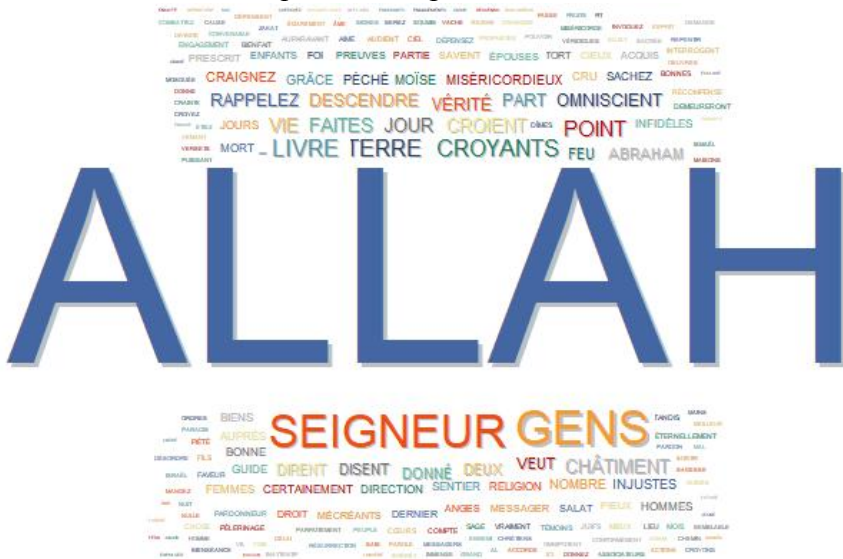
## 2.2 Fréquence des mots

### Fig.2.Nuage des mots de sourate Al Baqara

La fréquence des mots est à ne pas confondre avec la fréquence des codes. Par exemple, le mot Allah est apparu dans la traduction de Hamidulah en Français 296 fois alors que les versets codés Allah sont au nombre de 416. Ces codes sont attribués par le chercheur lorsque Allah est sous-entendu sans que le mot (Allah) lui-même n'apparaisse, ou lorsqu'un des Noms d'Allah est utilisé, comme Seigneur

**Il y a dans QDA miner deux options pour représenter la fréquence des mots :**

- Le tableau des fréquences simples



- Le nuage de mots

Nuage des mots de sourate Al Baqara : des nuages de mots interactifs

Une propriété utile de cette représentation, en plus d'être proportionnelle, est de retrouver l'ensemble des versets qui citent chacun des codes. Par exemple en cliquant sur le mot homme et en

choisissant chercher, le logiciel nous restitue les versets correspondants

Signalons ici un problème de cohérence ou de consistance : alors que le tableau des fréquences donnait 11 comme fréquences pour hommes et femmes, le nuage donne 10 pour les hommes. Nous avons utilisé la fonction statistique de Winword et si nous retrouvons bien 11 fois le mot hommes et 11 fois le mot femmes au pluriel, il y a 17 fois le mot homme et 14 fois le mot femme au singulier. Un travail est à faire sur le tableau car le nuage nous semble plus précis puisqu' il fournit les versets ou les mots hommes et femmes apparaissent.

L'utilisation des nuages de mots résume graphiquement en une seule image un texte qui peut prendre 25 pages, comme la traduction en Français par Hamidullah sous Word. C'est dire la puissance de cet outil qui aide à découvrir à travers les fréquences simples, les mots les plus proéminents.

Au vu du nuage, les principaux sujets abordés dans Al Baqara peuvent renvoyer aux thèmes suivants :

- La première des constatations est qu'ALLAH es plus grand que tout et ALLAH est au centre de tout
- Ses Livres, ses anges et ses Prophètes
- Création d'Adam
- Abraham
- Moïse
- Les Enfants d'Israël et leurs multiples désobéissances
- Les droits : législation pour la premier État musulman de Médine : mariage, divorce, héritages, contrats et autres

Il serait peu raisonnable de prétendre résumer ainsi sourate Al-Baqara qui est souvent présentée comme un condensé de tout le Saint-Coran. De la création d'Adam à la création de l'univers, les réponses aux questions essentielles peuvent y être trouvées.

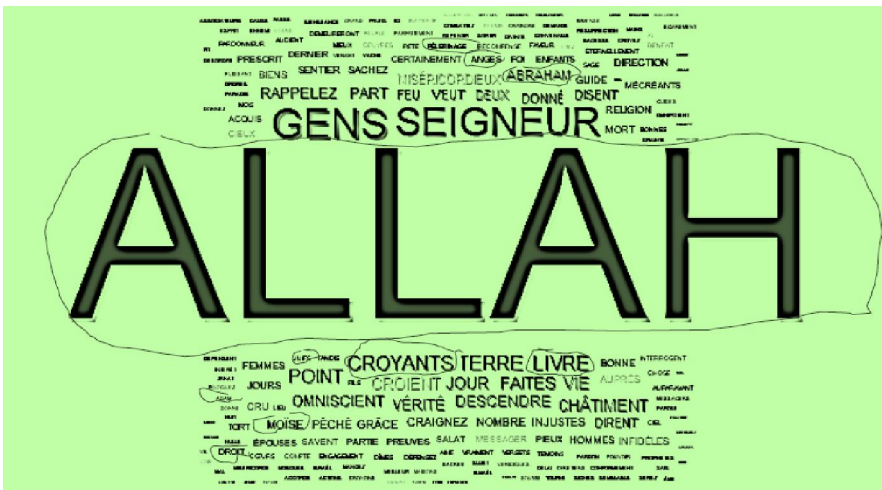
Notre but ici est purement pédagogique. Nous essayons de présenter un outil QDA miner pour explorer ses potentialités et ses limites dans

l'effort de rendre plus accessible le texte sacré traduit dans une autre langue.

Bien entendu si le chercheur n'a jamais lu sourate Al Baqara, il sera incapable de deviner les thèmes, simplement en regardant la représentation graphique. Cela est aussi valable pour un autre texte car le chercheur doit se familiariser avec le discours pour pouvoir établir des repères qui l'aideront à rattacher les mots hommes et femmes au mariage, au divorce, à l'héritage etc.

La méthode manuelle est utilisée ici en entourant les mots qui nous semblent importants dans la représentation graphique pour nous aider à résumer des thèmes importants.

### **Fig.3.Marquage des principaux mots qui pourraient renvoyer aux thèmes dans Al Baqara**



La puissance de cette représentation proportionnelle des fréquences de chaque mot est indiscutable. Devant le mot Allah, tous les autres paraissent minuscules.

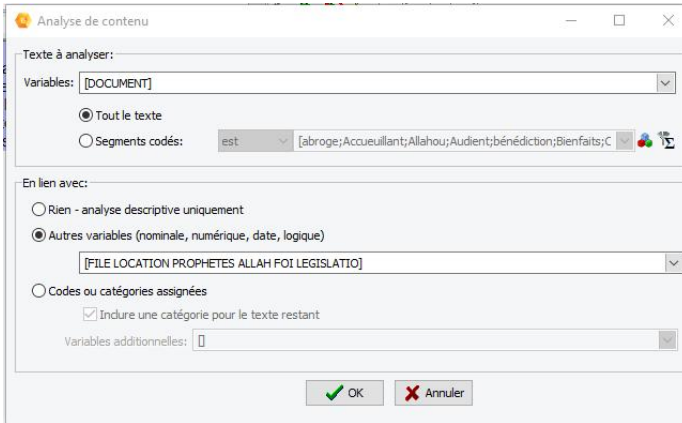
Une précision la catégorie Allah comprend les codes Allahou, Seigneur, etc.

L'essentiel des thèmes peuvent être induits à partir de ce nuage des fréquences des mots.

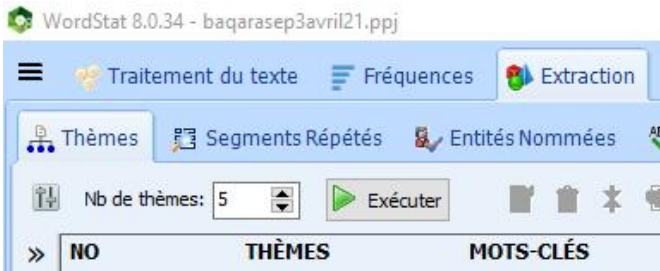
- Incontestablement l'unicité de Allah est le thème central puisque Son nom est mentionné plus de 400 fois dans cette sourate, qui ne compte que 286 versets, ce qui signifie que son nom Exalté revient plus d'une fois, en moyenne, par verset.
- La foi en Allah est nécessairement suivie de la reconnaissance de ses Livres, de ses anges et de ses Prophètes
- La création de Adam, son expulsion du paradis ainsi que la promesse d'Allah d'envoyer une guidance qui si elle est suivie assurera aux pieux la récompense et aux incroyants le châtement
- Si Abraham est le Prophète le plus cité dans la Sourate,
- Moïse suit en deuxième position
- Les multiples désobéissances des enfants d'Israël et leurs tergiversations à propos de la vache à sacrifier qui a donné son nom à la sourate
- La sourate, révélée à Médine, avec la constitution de la première société musulmane est ensuite venue avec la Loi qui doit régler les rapports entre les citoyens, mariage, divorce, héritage, le hajj et la omra, le prêt usuraire, les contrats de vente avec le verset le plus long, etc.

Bien entendu ces thèmes ne couvrent pas l'entièreté de la sourate mais seulement les thèmes les plus fréquents. D'autres classifications sont possibles. Avant de citer d'autres auteurs, il convient de présenter les thèmes identifiés par le QDA miner sous l'application WordStat.

Nous avons choisi de recourir à la fonction variable et d'utiliser les variables Allah (les multiples désignations), Prophètes (évoqués nommément ou inférés), et Foi (croyants, incroyants et hypocrites). Par exemple, pour renseigner la variable Prophètes, nous avons dû aller à travers tous les versets en signalant le nom du Prophète qui est cité dans le verset ou qui est concerné par ce verset. Dans le menu Analyse, il faut choisir Analyse de contenu



Une fois dans WordStat on doit choisir Extraction



En principe, le nombre de thèmes est à décider par le chercheur.

Dans notre cas, WordStat propose 9 thèmes et nous voulons simplement illustrer l'utilisation du programme sans discuter des limites. Ce point viendra plus tard.

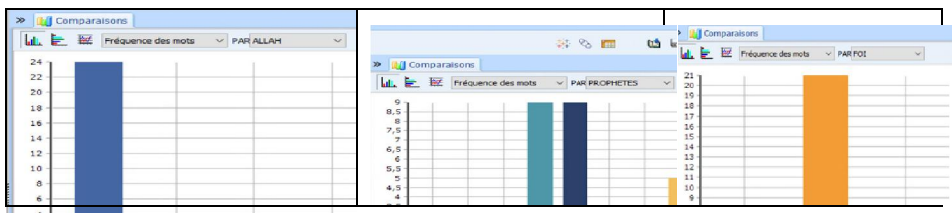
**Tableau 3.**Thèmes identifiés par WordStat

NO	THÈMES	MOTS-CLÉS
2	JESUS	PREUVES; TUÉS; JÉSUS; FILS; INFIDÈLES; GENS; IL; NOUS;


1	SOLAYMAN	DIABLES; PERSONNE; SOLAYMAN; BIEN; MÉCRÉANT; GENS; NOUS;
5	ABRAHAM	JE; MÉCRÉANT; MORT; VIE; GENS; SON; MOI; TU; ABRAHAM; ALLAH
3	TRIBUS	TRIBUS; ISAAC; ISMAËL; JACOB; ABRAHAM; JÉSUS; MOÏSE; PROPHÈTES; ALLAH;
4	ALLAH	BIEN; INFIDÈLES; IL; TU; ALLAH;
7	ALLAH ABRAHAM	MON; CRU; DIT; ABRAHAM; JE; ALLAH; LUI; SES; MOI;
8	ALLAH MOISE	MOÏSE; PROPHÈTES; SES; SON; IL; ALLAH; NOUS; CRU
6	ISAAC ISMAEL JACOB	DIVINITÉ; FILS; MORT; SES; MOI; IL; ISAAC; ISMAËL; JACOB; NOUS;
9	ALLAH	TU; ALLAH; MOI; IL; SON;

Ces thèmes générés par la machine sont modifiables : il suffit de cliquer sur le thème et de le modifier comme on le désire. Nous verrons plus loin comment les chercheurs ont conceptualisé des thèmes de manière beaucoup plus compréhensible.

**Fig.4.Fréquences simples pour les 3 variables Allah, Prophètes et Foi**



Dans le graphique, pour chaque thème, les Prophètes sont représentés par le nombre de mots et donc en cliquant sur le nom on peut obtenir les versets qui le concernent

Chacun des thèmes renvoie à Allah ou à un ou plusieurs de ses Prophètes, le thème 1 concerne le prophète Mohamed. 

En cliquant sur Montrer les mots clés on obtient les versets ayant reçu

le Code Muhammad



En cliquant sur un numéro de verset, QDA miner restitue en bas le texte du verset.

Recherche de mots-clés

Critère Résultats

CODE:

Cellules multilignes

Cas #	Texte	Apparue	Cas	FILE	LOCATION	PROPHETES	ALLAH	FOI	LEGISL
187	<a href="#">Cliquez pour voir l'item</a>	FIL; IL; ELLES;	Case #187 - baqara sep2	baqara sep2		Mohamed	Allahou	Croyants	retraite p
222	<a href="#">Cliquez pour voir l'item</a>	INTERROGENT;	Case #222 - baqara sep2	baqara sep2		Mohamed	Allahou	Croyants	menstrual
189	<a href="#">Cliquez pour voir l'item</a>	INTERROGENT;	Case #189 - baqara sep2	baqara sep2		Mohamed	Allahou	Croyants	pèlerin
286	<a href="#">Cliquez pour voir l'item</a>	IL; MAL;	Case #286 - baqara sep2	baqara sep2		Mohamed	Allahou	Croyants	charge int

189. Ils **interrogent** sur les nouvelles lunes - **Dis** : "Elles servent aux gens pour compter le temps, et aussi pour le Hajj [pèlerinage]. Et ce n'est pas un acte de bienfaisance que de rentrer chez vous par l'arrière des maisons . Mais la bonté pieuse consiste à craindre Allah. Entrer donc dans les maisons par leurs portes. Et craignez Allah afin que vous réussissiez !".

Bien entendu, il faut rappeler que la codification des thèmes par l'intelligence artificielle ne peut rivaliser avec l'intelligence humaine. Les limites de WordStat pour identifier des thèmes sont signalés dans le document cité en référence, en précisant que le programme ne pouvait prendre en charge plus de 2500 mots, or Al Baqara en compte 12000. Provalis travaille pour doubler cette capacité mais ce sera toujours insuffisant.

4. Ceux qui croient à ce qui t'a été descendu (révélé) et à ce qui a été descendu avant toi et qui croient fermement à la vie future.

5. Ceux-là sont sur le bon chemin de leur Seigneur, et ce sont eux qui réussissent (dans cette vie et dans la vie future).

Bien que le Prophète Mohamed ﷺ n'est pas cité nommément dans le verset 5, le verset 4 s'adresse clairement à lui, puisque la condition de la foi est de croire en Allah et en Mohamed comme Prophète. Le toi dans le verset 4 vise Mohamed ﷺ

Wordstat permet d'identifier les termes qui se répètent simultanément, que

l'on pourrait nommer co-occurrences mais que le logiciel présente comme phrases

### 2.3 Co-occurrence des codes

**COOCCURRENCE** - Une cooccurrence se produit chaque fois que deux codes apparaissent dans le même document (option à l'intérieur du cas). Vous pouvez également limiter la définition de la cooccurrence aux codes qui sont séparés par un nombre maximum de paragraphes (option Fenêtre de n paragraphes), ou aux codes qui en chevauchent d'autres (Option Chevauchement de segments).

Si nous nous concentrons sur les Messagers cités dans Al Baqara, nous pouvons trouver ceux qui sont cités dans le même verset, par exemple. On aurait pu choisir d'élargir le champ à 2 ou 3 versets, mais encore une fois notre but ici est purement démonstratif

Sur la figure 5, on remarque de prime abord 3 associations identifiées par des lignes :

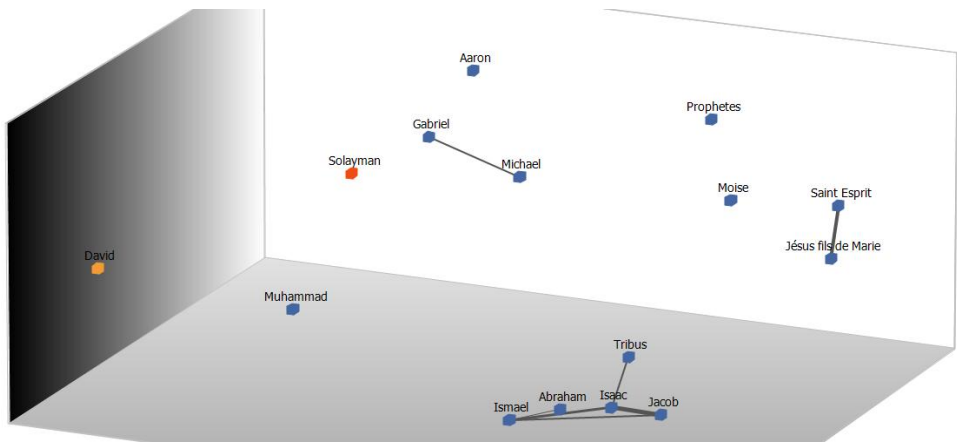
- Les anges Michael et Gabriel occupent une position à part des Envoyés humains.
- Abraham, le patriarche des Juifs, des Chrétiens et des Musulmans occupe une position centrale, non loin du trio Jacob, Isaac et Ismaël.
- Jésus fils-de-Marie et le Saint-Esprit.

Muhammad ﷺ n'est pas associé directement à un autre Messenger dans Al Baqara, ce qui signifie simplement que dans les



versets ou il est évoqué, on ne trouve pas de noms d'autres Envoyés. Alors que Jésus est souvent cité avec le Saint-Esprit, Abraham avec Ismaël, Jacob et Isaac.

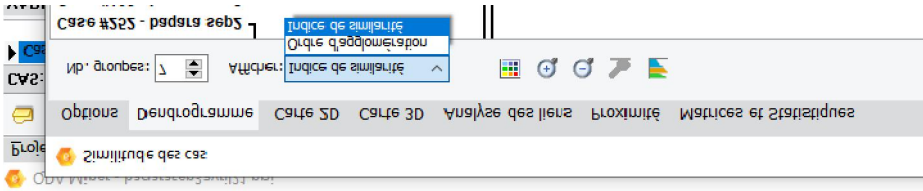
**Fig.5.Cooccurrences des noms des Prophètes dans Al Baqara**



### 2.3 Similitude des cas

QDA Miner utilise une méthode de classification hiérarchique pour créer des groupes. Le résultat est présenté sous forme de dendrogramme ou arborescence. Dans ce graphique, l'axe vertical se compose d'items et l'axe horizontal représente les groupes formés à chaque étape de la procédure de regroupement. Les codes qui tendent à apparaître ensemble sont combinés au début du processus d'analyse, alors que ceux qui sont indépendants les uns des autres ou ceux qui ne semblent pas apparaître ensemble tendent à être combinés à la fin du processus d'agglomération.

QDA miner offre le choix du nombre de groupes et de l'affichage par indice de similarité ou ordre d'agglomération



Rappelons que le cas correspond pour nous au verset. Si nous voulons savoir le nombre de fois qu'un Prophète est cité dans Al Baqara, QDA miner nous donnera les numéros des versets pour chacun des prophètes,

- Moïse, par exemple, apparaît dans les versets 51, 54, 55, 60, 61, 67, 92 et 108, 248, etc.
- Abraham apparaît dans les versets 124, 125, 126, 130, 135
- L'ange Gabriel est cité dans les versets 97 et 98 mais n'est identifié que par le verset 98
- Le Prophète Muhammad ﷺ est cité dans les versets 6, 80, 93, 94....

Tous les versets qui sont accolés ont en commun soit le nom d'un Prophète soit une autre similitude.

Par exemple en bas de la figure, on trouve accolés les versets 253 et 87

#### Verset 253

253. Parmi ces messagers, Nous avons favorisé certains par rapport à d'autres. Il en est à qui Allah a parlé ; et Il en a élevé d'autres en grade. A Jésus fils de Marie Nous avons apporté les preuves, et l'avons fortifié par le Saint-Esprit. Et si Allah avait voulu, les gens qui vinrent après eux ne se seraient pas entre-tués, après que les preuves leur furent venues ; mais ils se sont opposés : les uns restèrent croyant, les autres furent infidèles. Si Allah avait voulu, ils ne se seraient pas entre-tués ; mais Allah fait ce qu'il veut.

#### Verset 87

87. Certes, Nous avons donné le Livre à Moïse ; Nous avons envoyé après lui des prophètes successifs. Et Nous avons donné des preuves à Jésus fils de Marie, et

Nous l'avons renforcé du Saint-Esprit. Est-ce qu'à chaque fois, qu'un Messenger vous apportait des vérités contraires à vos souhaits vous vous enfliez d'orgueil ? Vous traitiez les uns d'imposteurs et vous tuiez les autres.

La similitude ici est constitué par la mention de Jésus fils-de-Marie et du Saint-Esprit alors que Jésus est cité dans d'autres versets.

**Fig.6. Citation des Prophètes selon les numéros des versets**



## 2.4 Séquence de codage

Les calculs de la séquence de codes cherchent à établir des régularités dans l'ordre d'apparition des codes. Allez dans la barre de menu, cliquez sur « Analyser » et sélectionnez « Séquences de codage »;

- Sélectionnez les codes initiaux et les deuxièmes codes que vous voulez; dans notre exemple nous avons simplement croisé certains noms d'Allah avec certaines des actions liées aux croyants, aux Enfants d'Israël et aux Chrétiens

- Paramétrez le calcul de la séquence avec les zones « Distance minimum » et « Distance maximale »; nous avons admis une proximité de 3 versets

Dans Expression de recherche, on doit choisir les codes en ligne et les codes en colonne. Une fois cela précisé, on clique sur chercher et la matrice de fréquences est automatiquement affichée.

Expression de recherche		Matrice de fréquences		Liste de fréquences		Résu		
Afficher: Fréquence de A suivi par B		Afficher: Couleurs & nombres						
A = indulgent B = aumone Fréq de A = 1 Fréq de B = 5 Fréq attendue = 0,0 B suit A = 1 (100,0%) A précède B = 1 (20,0%) % de séquences = 100,0% Valeur Z = 4,94 P = ,039		Chrétiens	Enfants d'Israël	Nature	aumone	Croyant	Dépense pieuse	gens du Livre
Agrément d'Allah						1		
grace d'Allah	4	1		1	1	1		
indulgent			1					
protecteur	1							

Pour chacun des cases de la matrice, on peut afficher les couleurs et les nombres, les couleurs seulement, ou les nombres seulement. Quant aux statistiques, nous avons le choix entre la fréquence de A suivi de

B, de B suivi de A, pourcentage de A suivi de B, pourcentage de B suivi de A, de la valeur Z, de la probabilité

Localisateur de séquences de code:1 résultats

Expression de recherche	Matrice de fréquences	Liste de fréquences	Résultats de recherche
Séquence: indulgent -> aumone			
Code:			
Cas	indulgent	aumone	
Case #263 - baqara sep2	263. Une parole agréable et un pardon valent mieux qu'une aumône suivie d'un tort. Allah n'a besoin de rien, et Il est indulgent.	263. Une parole agréable et un pardon valent mieux qu'une aumône suivie d'un tort. Allah n'a besoin de rien, et Il est	

Nous avons essayé de comprendre la séquence aumône-indulgence. En cliquant sur la cellule, on obtient le verset suivant :

**Une parole agréable et un pardon valent mieux qu'une aumône suivie d'un tort. Allah n'a besoin de rien, et Il est Indulgent.**

Le verset 263 associe bien aumône et indulgence mais pas nécessairement comme on pourrait s'y attendre : une parole agréable et un pardon valent mieux qu'une aumône suivie d'un tort. Ce qui signifie que toute aumône n'implique pas nécessairement l'Indulgence d'Allah, puisque si elle est suivie d'un tort elle a peu de chances d'être acceptée par Allah qui n'a besoin de rien.

Commentant ce verset, Ibn Kathir écrit que ceux qui donnent des aumônes et s'en vantent n'entreront pas au paradis.

Cette fonctionnalité de QDA miner semble particulièrement prometteuse, et pourrait donner des possibilités d'approfondissement de la compréhension du Saint-Coran en se référant à des commentateurs reconnus.

## 2.5 Codage par variables

La commande Codage par variables permet au chercheur d'examiner les modèles de codage par rapport aux variables numériques ou nominales. Les utilisateurs peuvent tester l'association entre les codes et les variables en choisissant entre onze tests de signification différents.

Les associations peuvent être graphiquement affichés sous forme de diagrammes à barres, de scatterplots, d'analyse de correspondance graphiques et des parcelles heatmap, qui peuvent tous être modifiés, imprimés, ou copié sur le presse-papiers.

Nous avons tenté de classer les versets de la sourate en 3 catégories, appelées dans QDA miner comme variable, ici la foi avec 3 catégories : croyants, incroyants et hypocrites. L'analyse de correspondances des fréquences des différents codes

REMARQUE : Les résultats sont limités aux trois premières dimensions, dont chaque combinaison peut être affichée comme un graphique 2D.

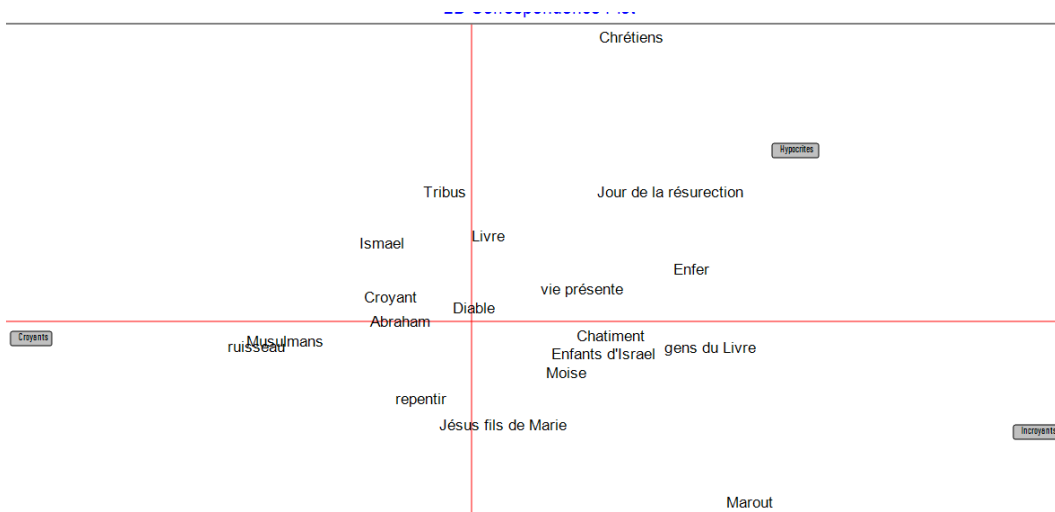
La figure 7 montre l'analyse par correspondance résulte d'un codage par la foi de la catégorie création. Il faut bien se rendre compte que cette classification dépend des choix pour le codage des versets et contient une bonne partie de subjectivité. Différents chercheurs pourraient arriver à des résultats différents. Nous ne cherchons ici qu'à montrer les fonctionnalités de QDA miner

- Associés aux hypocrites : le veau, l'enfer, les ténèbres, l'égarement, les Juifs et les Chrétiens
- Associés aux croyants : Hajj, Musulmans, Islam, soumis
- Associés aux incroyants : Les gens du livre, les Enfants d'Israël, intérêt usuraire, le châtement avilissant, Harout et Marout, le diable
- Le quatrième quadrant : on retrouve le repentir, les bonnes œuvres, la zakat, la Omra, le sentier D'Allah qui peuvent permettre aux incroyants de passer du côté des croyants.

De la même manière que les incroyants peuvent se repentir et passer du côté des croyants, les Croyants par leurs péchés

et leur égarement peuvent basculer dans l'hypocrisie.

**Fig.7. Analyse des correspondances selon la variable FOI (Croyants, Incroyants et Hypocrites)**



### 3. Analyse

Très justement, plusieurs auteurs soulignent le fait que la sourate s'ouvre comme une réponse à la supplication finale dans sourate al Fatiha : guide-nous vers le chemin droit. Al Baqara commence par 'Ce livre est un guide pour ceux qui craignent Allah'. Sourate Al Baqara aborde les thèmes les plus importants que l'on retrouvera avec des détails différents dans le reste du Saint-Coran

Voyons maintenant comment quelques auteurs ont identifié les principaux thèmes :

Raymond Farrin(FARRIN, 2010) a comparé les travaux de plusieurs auteurs qui ont identifié 6 thèmes (Islahi, Robinson et Zahniser) ou 5 thèmes (Smith). Il a quant à lui subdivisé la sourate en 9 parties successives avec des thèmes différents. Le tableau détaille les versets correspondant à chacun des thèmes

**Tableau 5 Classification des thèmes par versets(Farrin)**

Versets	Thèmes
A 1–20	Ce thème adresse essentiellement la subdivision en 3 groupes : croyants, incroyants et hypocrites
B 21–39	La création d’Adam, sa désobéissance et le pardon
C 40–103	Les enfants d’Israël et leurs multiples désobéissances
D 104–41	Les gens du livre Chrétiens et Juifs sont invités à accepter le Coran et l’Islam et à ne pas poser des questions au Prophète Mohamed Salla Allah Aleihioua
E 142–52	Changement de la Qibla et établissement des Musulmans comme communauté médiane distincte des Juifs et des Chrétiens
D 153–77	Autorisation est donnée aux Musulmans de pratiquer le rituel Safa et Marwa avec le récit de Hajer et de l’eau Zem Zem et précisions des interdits alimentaires
C178–253	Les règles individuelles et collectives qu’Allah a institué pour les Musulmans à Médine
B 254–84	S’adressant aux Croyants en les invitant à donner dans le sentier d’Allah et à respecter les engagements
A 285–86	Supplication à Allah.



Nevin Reda El-Tahrym(EL-TAHRYM, 2010) a soutenu une thèse de

Section	Versets	Thèmes	Character
§1.1	1-20	Classification de l'humanité en	Test
§1.2	21-29	Message direct à l'humanité	Instructions
§1.3	30-39	Histoire d'Adam et Eve	Histoire

doctorat à l'université de Toronto sur sourate Al Baqara et il a également proposé une classification des versets de la sourate en 9 thèmes principaux.

MOHAMAD IZZUDDIN RAHMAN, UN chercheur indonésien a choisi de classer les versets en 3 groupes (Iman, Ibadat, Akhlaq) et Nous sommes bien sur conscients que chacun de ces chercheurs avait ses propres objectifs et sa propre démarche méthodologique et c'est pourquoi il serait illusoire de s'attendre à retrouver exactement les mêmes résultats. Cependant si on compare ces auteurs et leurs thèmes à ceux identifiés par WordStat, force est de constater que l'identification des thèmes par le logiciel est beaucoup moins pertinente que les autres.

### **Tableau 6 Classification des thèmes par versets(Tahrym)**

§2.1	40-48	Instructions aux Enfants	Instructions
§2.2	49-74	Interactions passées entre les	Story
§2.3	75-123	Interaction avec la	Test
§3.1	124-	Origines abrahamiques	Histoire
§3.2	152-	Législation pour la nouvelle	Instructions
§3.3	243-	Test de la Foi	Test

#### 4. Discussion

Une manière de concilier les résultats des différentes approches pour caractériser les principaux thèmes de Al Baqara serait de comparer les poids respectifs accordés à chacun des thèmes en termes de nombre de versets, de codes ou de mots.

**Tableau 7 Pourcentage des versets selon lesthèmes (Farrin)**

Versets	Thèmes	%	Verset	Thèmes	%
A 1–20	Subdivision en	7%	A 1–20	Classification	7%
B 21–39	La création	6%	21-29	Message	3%
C 40–103	Les enfants	22%	30-39	Histoire	3%
D 104–41	Les gens du	22%	40-48	Instructions	3%
E 142–52	Changement	4%	49-74	Les Enfants	9%
D 153–77	Autorisation	8%	75-	La	17
C 178–253	Les règles d'Allah	26%	124-151	Origines abrahamique	9%
B 254–84	Croyants	11%	152-242	Législation pour la nouvelle Nation	31 %
A 285–86	Supplication à Allah.	1%	243-286	Test de la Foi	15 %

Malgré certaines différences, il y a un grand accord sur les thèmes eux-mêmes et leur formulation.

- On signale surtout le poids des versets consacrés à la législation pour la nouvelle communauté de Médine (45% chez Mattin et 48% chez Tahrymi). Autrement presque la moitié de sourate Al Baqara est dédiée aux règles individuelles et collectives pour les Musulmans.
- La subdivision de la population de Médine en 3 groupes est évoquée avec la même proportion de versets chez les deux auteurs (7%) : croyants, incroyants et hypocrites
- On note une différence pour les Enfants d'Israël (22%) chez Mattin et (12%) chez Tahrym.

On remarque deux lacunes chez les deux auteurs :

- Clairement Allah est central dans cette sourate et aucun des auteurs ne le souligne
- Muhammad ﷺ ans jamais être cité nommément est également le plus présent<sup>1</sup>des prophètes sans que ces deux auteurs n'aient pu l'identifier

On remarque que la subdivision par versets est trop simpliste, Ainsi le Prophète Muhammad ﷺ apparaît tout au long de la sourate sans que son nom ne soit jamais explicite : les versets

---

1

**252. Voilà les versets d'Allah, que Nous te (Muhammad)  
récitons avec la vérité. Et TU ES, certes parmi les  
Envoyés.**



**Tableau 8 Pourcentage des versets et pourcentage des codes**

Catégorie	Code	Fréquence	% Codes	Cas	% Cas
\	Allah	441	2370,0%	200	6990,0%
Création	Enfants d'Israël	58	310,0%	58	2030,0%
Croyants	Musulmans	41	220,0%	41	1430,0%
Envoyés	Muhammad	33	180,0%	33	1150,0%
Envoyés	Livre	25	130,0%	23	800,0%
Envoyés	Abraham	16	90,0%	12	420,0%
Envoyés	Moïse	12	60,0%	12	420,0%

Il faut signaler ici une erreur dans le calcul de WordStat : pourcentages sont multipliés par 100

Autre anomalie : impossibilité d'avoir plusieurs catégories pour la variable : on ne peut avoir qu'un prophète alors que plusieurs peuvent être cités dans le même

Il vaut mieux donc travailler sur le nuage de codes plutôt que sur le nuage de mots.

## 5. Conclusion

Il est beaucoup plus intéressant d'utiliser le nuage de codes pour identifier les thèmes que se contenter des résultats de WordStat

Le compteur semble poser un problème : le nombre total de mots de la sourate est 12709 pour la version française de QDA miner et 12733 pour la version anglaise. Selon Word, le compte serait de 12560.




Les aspects graphiques aident énormément à résumer l'information plus que les mots ou les chiffres

Nous recommandons donc l'utilisation de QDA miner mais avec réserve sur certaines de ses fonctionnalités

- Anomalie dans le calcul de WordStat : pourcentages sont multipliés par 100
- Autre anomalie : impossibilité d'avoir plusieurs catégories pour la variable : on ne peut avoir qu'un prophète alors que plusieurs peuvent être cités dans le même verset.
- Bien entendu, il faut rappeler que la codification des thèmes par l'intelligence artificielle ne peut rivaliser avec l'intelligence humaine. Les limites de WordStat pour identifier des thèmes sont signalés dans le document cité en référence, en précisant que le programme ne pouvait prendre en charge plus de 2500 mots, or Al Baqara en compte 12000. Provalis travaille pour doubler cette capacité mais ce sera toujours insuffisant.
- Le fait que les thèmes peuvent changer pose un problème et il ne faut guère accorder de crédit aux thèmes identifiés par WordStat à ce stade du développement du programme. Lors du webinaire de Provalis, j'ai posé une question sur le fait que WordStat pouvait donner des résultats différents pour les thèmes et la réponse qui fut donnée est la suivante : NMF LIKE LDA ARE PROBABILISTIC, SOWILLNEVERGIVE THE EXACT SAMERESULT. BUT TOPIC MODELIND FROM FA WILL GIVE SAME RESULTS (BASED ON DETERMINISTIC

APPROACH) LATENT DIRICHLET ALLOCATION - A TOPIC MODELING TECHNIQUE.

- En termes d'identification des thèmes, nous pensons que le découpage linéaire en versets a montré des limites puisque cette manière de découper le texte a laissé échapper deux thèmes primordiaux la Grandeur d'Allah et son interlocuteur Muhammad  qui sont présents tout au long de la sourate et non concentrés dans un nombre limité de versets.

Il reste à signaler que la version arabe de QDA miner serait un véritable atout pour les chercheurs d'expression arabe, d'autant plus que les sourates du Saint-Coran ont été révélées dans cette langue. Travailler sur une traduction n'égalera jamais la version originale. Qu'il suffise de signaler que la version arabe compte environ 6000 mots alors que la traduction de HamiduAllah pratiquement le double.<sup>2</sup>

Une précaution méthodologique en ce qui concerne l'interprétation des résultats de QDA miner : le chercheur doit toujours se méfier de ses propres interprétations et confronter sa compréhension à celle des érudits établis et reconnus par les Musulmans, sinon le risque est grand de se voir accuser de Bidaa ou Innovation dans le mauvais sens du terme.

## Références

- (n.d.). Retrieved April 12, 2021, from <https://qda-miner-lite.software.informer.com/4.0/>: <https://qda-miner-lite.software.informer.com/4.0/>
- BARRY R. & MAAS, S. (2007, February). Review of QDA miner. *Field Methods*, 19(1), 87-108.
- COLIN, SIMON. (2010). *TUTORIEL PRISE EN MAIN DE QDA MINER POUR L'ANALYSE DES DONNÉES QUALITATIVES*. Université de Montréal, Montréal.

- EL-TAHRYM, N. R. (2010). *TEXTUAL INTEGRITY AND COHERENCE IN THE QUR'AN: REPETITION AND NARRATIVE STRUCTURE IN SURAT AL-BAQARA*. University of Toronto, DOCTOR OF PHILOSOPHY NEAR AND MIDDLE EASTERN CIVILIZATIONS, Toronto.
- FARRIN, R. K. (2010, January). SURAT\_AL\_BAQARAH\_A\_STRUCTURAL\_ANALYSIS. *THE MUSLIM WORLD*, 100. Retrieved from <https://www.academia.edu/8642515>
- HAMIDULLAH, M. (1985). *LE SAINT CORAN*. MARYLAND, U.S.A.: AMAZONA CORPORATION.
- OUDA, K. (2015). *Quran Analysis: a semantic search and intelligence system for the Quran*. These de Msc, University of Leeds, Leeds.
- (2018). *QDA miner, Logiciel d'analyse qualitative des données, Manuel de l'utilisateur*. Recherches Provalis, Montréal.
- RAHMAN, M. I. (2018). Comparative Analysis for Topic Classification in Juz Al-Baqarah. *Indonesian Elec Eng & Comp Sci*, 12(1), 406-411.
- VOYNNET-FOURBOUL. (2012). *FORMATION QDA MINER PARIS*. UNIVERSITE PANTHEON-CIFFOP, Paris.
- (2015). *WORDSTAT 7, USER'S GUIDE*,. RECHERCHE PROVALIS, MONTREAL.