



REPERER DES DIMENSIONS DANS LES REX : UTILISATION DU TAL EN MILIEU MEDICAL

S. Lagarde, C. Raynal, A. Urieli

► To cite this version:

S. Lagarde, C. Raynal, A. Urieli. REPERER DES DIMENSIONS DANS LES REX : UTILISATION DU TAL EN MILIEU MEDICAL. Congrès Lambda Mu 21 “ Maîtrise des risques et transformation numérique : opportunités et menaces ”, Oct 2018, Reims, France. hal-02074925

HAL Id: hal-02074925

<https://hal.science/hal-02074925>

Submitted on 21 Mar 2019

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.

REPERER DES DIMENSIONS DANS LES REX : UTILISATION DU TAL EN MILIEU MEDICAL

LOCATING DIMENSIONS IN FEEDBACK REPORTS: USING NLP IN THE MEDICAL DOMAIN

Lagarde S.
UGECAM ALPC et Centre
8 rue de Limoges
87 430 Verneuil sur Vienne

Raynal C.
Safety Data – CFH
13 rue Temponières
31 000 Toulouse

Urieli A.
Joliciel Informatique
2 avenue du Cardé
09000 Foix

Résumé

Grâce aux DEI, les agents des établissements de santé gérés par l'UGECAM ALPC et Centre signalent tout événement non attendu rencontré. Ces retours d'expérience sont utilisés par les responsables qualité pour analyser les situations problématiques et/ou à risque et améliorer la sécurité et la qualité des soins et services dispensés. Afin de les aider dans leur travail d'analyse, la web application *PLUS* d'aide à l'exploitation des données textuelles a été mise en place ; elle contient un module de création de dimensions qui permet de repérer les DEI porteuses d'une problématique transverse, difficilement repérables par des recherches classiques et non identifiables via les métadonnées existantes. Nous rendons ici compte de la plus-value de cette démarche et des résultats obtenus sur le thème de la violence.

Summary

Employees in medical establishments managed by UGECAM ALPC et Centre use undesirable event reports to signal the occurrence of unexpected events. This feedback is used by the quality managers to analyse problematic and/or risky situations and improve the safety and quality of the treatments and services provided. The reports are analysed with the help of the textual data mining web application *PLUS*. *PLUS* contains a dimension training module, which makes it possible to locate reports referring to cross-domain problems which are difficult to locate via conventional key-word searches and unidentifiable using exiting metadata. We describe the added value of this process and the results obtained for the study of violence.

Introduction

Grâce aux Déclarations d'Événement Indésirable (DEI), les agents des 13 établissements de santé gérés par l'UGECAM ALPC et Centre signalent tout événement non attendu ayant lieu au cours de leur activité. Les DEI constituent la pierre angulaire de la gestion des risques et de la démarche qualité au sein des établissements de santé dont l'UGECAM a la charge. En effet, c'est notamment par leur analyse que des Plans d'Amélioration de la Qualité (PAQ) sont définis et mis en œuvre.

Lors de la déclaration, les agents sont invités à renseigner un certain nombre de champs, comme la date et la tranche horaire de l'événement ou encore le pôle sur lequel il s'est produit. Ces champs qui contiennent des métadonnées sont complétés par une description totalement libre (*i.e.* non orientée par des questions et/ou la sélection de mots-clés par exemple) rédigée en langue naturelle par l'agent qui déclare. En 2017, environ 160 événements ont été déclarés chaque mois et la base, qui contient les DEI de janvier 2010 à février 2018, est à ce jour constituée d'un peu plus de 12 250 déclarations.

Grâce à l'acculturation sécurité et qualité ainsi qu'à la facilitation du processus de déclaration, on voit le volume de ces retours d'expérience croître. En effet, en 2011 on comptait en moyenne 70 DEI par mois, chiffre qui a atteint 150 DEI mensuelles en 2013 ; puis une décroissance a été observée pour atteindre 112 déclarations mensuelles en moyenne en 2015, pour repartir à la hausse et atteindre 160 DEI en 2017 suite à une campagne d'information faite auprès des agents et 200 sur les premiers mois de 2018.

Si on peut se réjouir de cette croissance globale dans la mesure où elle permet de dresser un panorama plus exhaustif et plus réaliste de la vie des établissements, elle confronte les Responsables d'Amélioration de la Qualité (RAQ) à un problème majeur, à savoir le temps nécessaire pour analyser ces déclarations. En effet, il revient aux RAQ de traiter les DEI afin d'établir leur lien avec le processus et plus précisément l'activité en jeu dans chacune d'elles (par exemple l'accueil, la prise en charge médicamenteuse ou encore le pilotage de l'établissement), d'identifier les

causes sous-jacentes¹ ainsi que les risques qui y sont liés², et enfin, une fois ces catégorisations effectuées, de définir les PAQ pertinents.

Afin d'aider les RAQ dans leur travail d'analyse, l'application d'aide à l'exploitation des données textuelles *PLUS* a été mise en place en juin 2015³ : elle permet non seulement de faire des recherches puissantes dans les textes des DEI⁴ mais également de catégoriser

¹ Le modèle utilisé est inspiré du diagramme d'Ishikawa ou 5M (Matière, Matériel, Méthode, Main d'œuvre et Milieu).

² Les processus couvrent l'ensemble des missions d'un établissement, qu'elles concernent le management (ex : « Gérer les ressources humaines »), les fonctions supports (ex : « Gérer la fonction linge ») ou le cœur de métier (ex : « Développer le projet d'accompagnement et de soins personnalisé »). Chaque processus est détaillé en activités (ex : sur le processus ressources humaines, « Gérer la performance, les compétences, les carrières et les évolutions salariales, les entretiens annuels »). De même les activités comprennent plusieurs tâches (ex sur l'activité précédemment citée, l'évaluation des compétences professionnelles). Enfin, chaque tâche représente un risque (ex : « L'évaluation des compétences professionnelles n'est pas assurée ») ; en effet les risques sont des tâches non ou mal réalisées.

³ *PLUS*, pour *Processing Language Upgrades Safety*, est une application web développée par Safety Data - CFH dédiée à l'exploration et l'exploitation des bases de données textuelles, de REX notamment.

⁴ Deux outils complémentaires, et qui s'appuient tous les deux sur les textes analysés linguistiquement (cf. plus loin, section « Traitements linguistiques des DEI »), permettent de réaliser des recherches fines et d'obtenir des résultats complets rapidement : le moteur de recherche permet de spécifier des requêtes via des mots ou expressions clés tandis que le module d'analyse de similarité textuelle permet de compléter les résultats en rapprochant des documents grâce à leur proximité linguistique.

automatiquement celles-ci afin d'identifier l'activité en jeu, le risque, sa ou ses causes ainsi que sa gravité⁵.

Si ces catégorisations sont cruciales pour analyser individuellement les DEI et permettre leur analyse collective grâce au tri ainsi effectué, il n'en demeure pas moins que ce tri est fait selon les taxonomies disponibles et les valeurs prédéfinies, autrement dit la catégorisation porte en elle une limitation liée au fait qu'on analyse les événements en classes nécessairement pré-identifiées. Il est certes possible de créer de nouvelles catégories mais cela doit être fait avec parcimonie dans un souci de cohérence et de pérennité dans l'analyse, et tout ajout (ou suppression) d'une catégorie nécessite de revoir l'intégralité de la catégorisation déjà appliquée afin de prendre en compte l'ajustement globalement.

Par ailleurs, il arrive que l'on souhaite repérer les DEI relatives à un phénomène particulier mais non identifié dans les catégories existantes parce que transverse à ces classes prédéfinies soit qu'il est émergent soit qu'on veuille suivre son évolution de manière prospective. Ainsi, les experts de l'UGECAM peuvent vouloir chercher des DEI qui partagent un thème, appelé « dimension », sans pour autant partager la même activité et/ou être liées directement à la même cause. De la même manière, les experts peuvent chercher à savoir s'il y a un contexte qui favorise l'apparition de plusieurs événements, indépendamment du risque précis associé à chacun d'entre eux afin de définir les plans d'action qui, en prenant en charge la dimension commune à ces événements, vont permettre de prévenir les risques associés.

Pour répondre à la nécessité d'identifier les DEI relatives aux différentes dimensions pertinentes pour les experts UGECAM et absents dans les catégorisations existantes, et pour aller ainsi plus loin dans l'aide à l'analyse des données textuelles, une approche complémentaire à ces catégorisations est préconisée : l'analyse en dimensions, selon la technique introduite dans Tanguy *et al*, 2015. Cette technique est depuis passée de la phase expérimentale à la phase opérationnelle : un module de création et de gestion de dimensions est maintenant disponible dans l'application *PLUS*, qui permet aux utilisateurs de définir un thème – appelé dimension, ou encore problématique transverse – afin de pouvoir identifier rapidement, *via* une métrique, les documents qui en sont porteurs. Il s'agit ici de rendre compte de la méthodologie et de la plus-value d'une telle fonctionnalité ainsi que des premiers résultats obtenus depuis sa mise en service fin 2017.

Le premier aspect méthodologique qu'il nous semble crucial de présenter concerne la démarche des experts de l'UGECAM et leurs motivations à utiliser les dimensions. Nous détaillerons ensuite les aspects techniques relatifs à la mise en place fonctionnelle de la création et la gestion de dimensions dans l'application *PLUS*, puis présenterons l'évaluation des résultats obtenus avant de proposer quelques éléments de conclusion.

Contexte et enjeu experts

Dans le contexte de l'UGECAM, la création de dimensions est justifiée dans principalement deux cas :

- Vérifier le respect de la législation
- Confirmer l'émergence d'une problématique

La vérification du respect de la législation est nécessaire lorsqu'une nouvelle réglementation entre en vigueur comme avec un texte plus ancien. Dans le cas d'une nouvelle réglementation, l'établissement doit inscrire ce

point dans son programme d'amélioration de la qualité et de la sécurité de soins. Il est fort probable que l'existant ou le contexte ne soit pas encore cerné. Généralement, l'établissement réalise une analyse initiale pour définir le périmètre de travail ; la dimension apportera une aide sur ce point. Dans le cas où le texte est plus ancien, il s'agit de s'assurer du respect de la réglementation dans les pratiques. L'établissement dispose d'une organisation définie sur le thème concerné mais pour lequel une évaluation est nécessaire. On prendra en exemple la lutte contre la maltraitance qui est une priorité depuis plus de 10 ans dans les établissements de santé. Ce thème a fait l'objet de recommandations formalisées puis d'une première évaluation en fin 2015. En 2018, une nouvelle évaluation est à mener. La dimension va permettre d'identifier les événements porteurs, premiers éléments pour cette évaluation.

Dans le cas où il ne s'agit plus de faire face à une nouvelle réglementation mais de confirmer l'émergence d'une problématique, l'établissement identifie celle-ci à partir d'un point d'entrée qui peut être d'origines diverses :

- une déclaration d'évènement indésirable,
- un sujet abordé au cours d'une réunion,
- un courrier de plainte,
- une réclamation orale, etc.

Cette identification se trouve isolée et ne peut constituer à elle seule une raison de modification de l'organisation. Il est nécessaire de faire une analyse de la situation globale pour établir le besoin d'évolution. La dimension va permettre de repérer l'ensemble des déclarations d'évènement indésirable déjà produites sur ce thème. Les données repérées seront la base d'une analyse pour identifier les actions d'améliorations pertinentes.

On peut par exemple noter que le respect des droits des patients est une préoccupation constante des établissements mais l'évaluation de ce respect est extrêmement difficile car il touche toutes les activités de l'établissement. Aujourd'hui cette évaluation est soit ciblée sur un sujet (comme l'information à transmettre sur les directives anticipées⁶) soit réalisée à partir du ressenti exprimé par les patients au travers d'enquêtes de satisfaction. Une dimension « Droits des patients » créée à partir des DEI va permettre d'identifier les problèmes de non-respect des droits quel que soit le sujet en cause et en évitant le biais lié au ressenti patient.

Il est enfin intéressant de noter que les éléments que l'on souhaite surveiller peuvent être disruptifs : des problématiques émergent lors de la mise en place de nouvelles organisations. On observe par exemple qu'afin d'améliorer la participation des patients, il leur est désormais remis un planning de rendez-vous et d'activités pendant leur séjour hospitalier. Ce planning leur permet plus d'autonomie mais certains patients ne respectent pas les horaires par oubli, négligence ou incompréhension. La dimension a ici pour objectif d'identifier les impacts négatifs sur les activités de l'établissement afin d'évaluer ces conséquences et d'ajuster au mieux la remise du planning en la personnalisant.

Mise en place technique

Les enjeux pour les experts ayant été présentés, nous décrivons les aspects techniques relatifs à la mise en place du module de création et de gestion de dimensions, à savoir les traitements linguistiques automatiques appliqués au contenu textuel des DEI ainsi que la méthodologie propre aux dimensions, avant de détailler les résultats relatifs à la mise en place d'une dimension.

⁵ Le module de catégorisation automatique disponible dans *PLUS* ne sera pas décrit plus en détail ici ; nous renvoyons à l'article proposé dans le cadre du λμ 21 « Utilisation du TAL pour la consolidation de données de fiabilité issues des REX » (Gaucher *et al.*, 2018).

⁶ Toute personne majeure peut rédiger des directives anticipées pour exprimer ses souhaits concernant sa fin de vie et une information doit être faite au patient dans ce sens par les professionnels de santé.

1 Traitements linguistiques des DEI

Comme précisé en introduction, les DEI contiennent deux types d'informations lorsqu'elles sont envoyées vers PLUS : des métadonnées factuelles (établissement, date, etc.) et une description en langue naturelle. Ces deux types de données sont traités indépendamment l'un de l'autre, les informations textuelles nécessitant des traitements spécifiques.

En effet, la description textuelle de toute DEI intégrée à PLUS fait l'objet d'une analyse linguistique automatique qui consiste en plusieurs étapes. Ainsi, une fois la langue identifiée⁷, la suite de traitements linguistiques Talismane⁸ effectue un certain nombre de pré-traitements afin :

- de corriger des fautes d'orthographe récurrentes (« *traitemnt* » sera corrigé en « *traitement* »),
- d'uniformiser les différentes formes d'un même terme telles que les abréviations (« *ttt* » est équivalent à « *traitement* »),
- de rapprocher les acronymes de leur version longue (« *ARS* » et « *Agence Régionale de Santé* »),
- d'identifier des « entités nommées », à savoir des mots ou groupes de mots qui réfèrent à une entité unique : on identifie ainsi classiquement les noms de personnes, de lieux ou d'organisation comme celui des établissements ou les noms de médicaments.

Ces pré-traitements s'appuient à la fois sur une étude linguistique des données analysées mais aussi sur des ressources propres à l'UGECAM : en effet, si « *abs* » est l'abréviation de « *absent* » dans le présent contexte, il en va différemment dans le milieu automobile où l'on reconnaît l'acronyme du système anti-blocage des roues (« *Antilock Brack System* »). Afin d'analyser au mieux les DEI, il est par conséquent crucial de prendre en compte les spécificités linguistiques propres au domaine médical qu'est celui de l'UGECAM.

Une fois ces différents pré-traitements réalisés, un module Snowball⁹ est utilisé pour « stemmer » les textes : chaque terme de chaque phrase est associé à sa racine (son *stem*). Ainsi des termes comme « *agression* » et « *agresser* » ont la même racine « *agress* », ce qui permet par conséquent de retrouver les textes qui contiennent aussi bien l'un que l'autre lorsque l'on recherche l'un ou l'autre.

L'ensemble des traitements automatiquement réalisés sur les données textuelles permet de transformer ces informations que l'on peut qualifier d'initialement « non structurées » (comparativement à des métadonnées) en objets structurés dans lesquels la variation linguistique est maîtrisée, permettant ainsi de réaliser des traitements statistiques pertinents et efficaces. Ce ne sont plus les mots tels qu'ils sont spécifiquement rédigés qui sont considérés par la suite mais leur représentation, les « traits » correspondants. En d'autres termes, ce sont sur ces traits et les ensembles qu'ils constituent que sont basés tous les traitements ultérieurs : ceux permettant de faire des recherches comme ceux utilisés pour l'analyse de similarité, ceux pour la catégorisation automatique comme ceux pour la création de dimensions.

2 Objectif des dimensions

Afin d'illustrer la méthodologie utilisée pour créer des dimensions, nous prendrons l'exemple de la violence, plus

large mais directement liée à la problématique de la maltraitance évoquée plus haut. En effet, celle-ci est transverse pour l'UGECAM dans le sens où elle peut apparaître aussi bien dans une activité de soin que dans le cadre de la distribution des repas, et peut concerner aussi bien les patients ou leur entourage que les professionnels qui interviennent dans les établissements de santé. Par ailleurs, la violence n'est pas présente en tant que telle dans la liste des risques qui sont organisés autour des activités.

Par conséquent, jusqu'à présent, la seule façon d'identifier les DEI dans lesquelles la violence est présente consistait à utiliser la recherche classique, par mots-clés. Mais celle-ci ne suffit pas toujours car, aussi puissante soit-elle, elle requiert la définition d'expressions-clés dont il est difficile d'être sûr de l'exhaustivité comme de la non-ambiguïté.

Ainsi, en cherchant par exemple le terme « *violent* », on trouvera bien les DEI mentionnant les patients violents (Figure 1) mais cela ne suffira pas à repérer toutes les DEI évoquant l'agressivité (physique ou verbale) qui est souvent exprimée autrement. On pourra compléter en cherchant le terme « *coup* » qui permettra de trouver les DEI mentionnant un « *coup de poing* » ou un « *coup de tête* », mais on note immédiatement que ces expressions peuvent aussi bien être le signe de violence physique que d'impulsivité totalement dénuée d'agressivité, et on trouvera par ailleurs des DEI mentionnant un « *coup de main* » absolument non ambigu et non pertinent pour la violence.

Ainsi, les résultats pertinents se trouvent mêlés à la fois à :

- du « bruit », à savoir de résultats répondant bien aux requêtes spécifiées mais non pertinents par rapport à ce que l'on cherche,
- et du « silence », plus difficile à identifier en raison même de sa nature, à savoir les DEI évoquant la violence sans en faire mention via les mots et expressions clés recherchés. Il sera par exemple délicat de trouver une DEI mentionnant simplement que « *le patient de la chambre 10 a traité son voisin de tous les noms* » alors qu'aucun des termes clés qui viennent spontanément à l'esprit ne sont présents (« *violence* », « *agressivité* », « *coup* », « *insulte* », etc.) d'une part et qu'à l'inverse, le terme « *traité* » est lui particulièrement ambigu dans le contexte médical dans lequel nous sommes d'autre part.

L'objectif, dans un exemple comme celui de la violence, est d'identifier efficacement les DEI qui se rapportent à cette problématique, qu'elle soit explicitée de façon évidente ou évoquée de manière plus implicite. Pratiquement, il s'agit donc de proposer un moyen de compléter efficacement la recherche classique par mots-clés afin de réduire le bruit présent dans les résultats (les résultats non pertinents) tout en évitant le plus possible le silence (les résultats pertinents non identifiés).

3 Création de dimensions

Afin de répondre aux enjeux et objectifs décrits ci-dessus, un module fonctionnel existe dans l'application PLUS permettant à un utilisateur de créer et de gérer des dimensions. Une dimension va alors correspondre, une fois créée, à une nouvelle métadonnée numérique attribuée à chaque DEI de la base de données ; métadonnée dont la valeur est comprise entre 0 et 1 qui permet de mesurer la présence de la dimension dans un document : plus la valeur de la dimension se rapproche de 1, plus la DEI est porteuse de la dimension en question, tandis que plus le poids tend vers 0, moins celle-ci est jugée présente dans le document.

Pratiquement, pour une dimension donnée choisie par l'utilisateur, celui-ci est invité à constituer un « ensemble positif » de DEI, ensemble qui regroupe des DEI dans lesquelles il juge la dimension présente, explicitement ou implicitement. Les DEI peuvent être sélectionnées grâce au moteur de recherche : en spécifiant

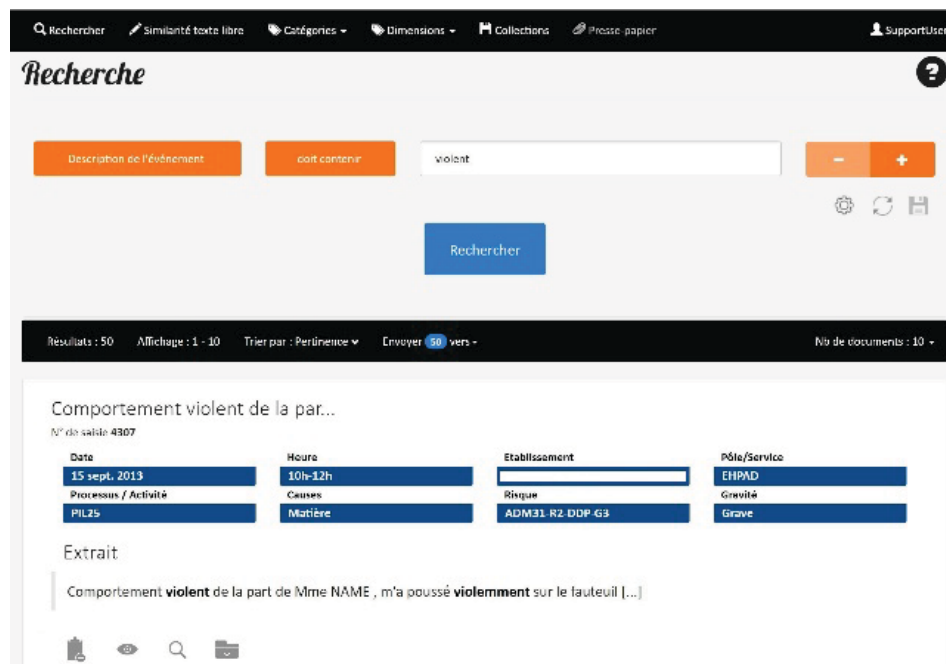
⁷ Cette étape est relativement triviale dans le cas des DEI car elles sont toutes, sans exception, rédigées en français. On fera néanmoins remarquer que des bases de données textuelles françaises peuvent comporter des phrases voire des documents entiers en anglais ; on citera le cas des bases de données aériennes comme les rapports de vol (*Air Safety Reports*) par exemple.

⁸ Cf. (Urieli, 2013).

⁹ Cf. (Porter, 2001).

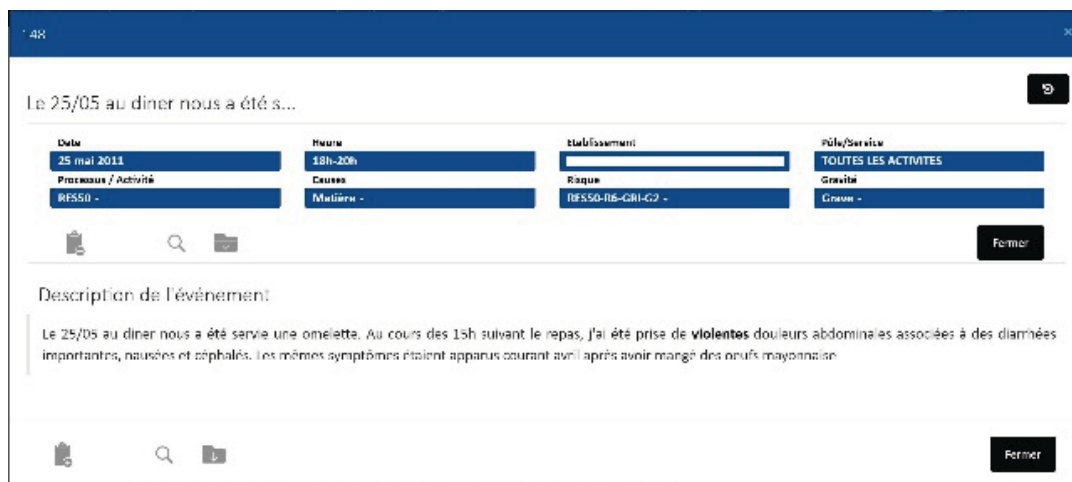
leur identifiant si elles sont connues de l'expert ou en effectuant des requêtes par mots-clés afin de regrouper les DEI les plus représentatives de la dimension qu'on cherche à modéliser. C'est ce qu'illustre la Figure 1 dans laquelle on voit l'un des résultats de la recherche du mot « violent » : DEI pertinente qu'on mettra dans l'ensemble positif.

On doit toutefois faire preuve de vigilance car, comme évoqué précédemment, les résultats peuvent contenir des DEI non pertinentes comme on le voit dans la Figure 2 avec « j'ai été prise de violentes douleurs ».



The screenshot shows a search interface titled 'Recherche'. At the top, there are navigation links: 'Rechercher', 'Similitude texte libre', 'Catégories', 'Dimensions', 'Collections', 'Presse papier', and 'Support User'. Below the title, there are two orange buttons: 'Description de l'événement' and 'contenir', followed by a text input field containing the word 'violent'. To the right of the input field are minus and plus icons. Below these is a blue 'Rechercher' button. Underneath the search bar, there is a status bar showing 'Résultats : 50', 'Affichage : 1 - 10', 'Trier par : Pertinence', 'Envoyer 50 vers', and 'Nb de documents : 10'. The main content area displays a result titled 'Comportement violent de la par...' with a sub-header 'N° de santé 4307'. Below this, there are several filters: 'Date' (15 sept. 2013), 'Heure' (10h-12h), 'Etablissement' (empty), 'Pôle/Service' (EHPAD), 'Processus / Activité' (PIL25), 'Causes' (empty), 'Risque' (ADM31 R2 DDP G3), 'Gravité' (Grave), and 'Matière' (empty). An 'Extrait' section shows a snippet of text: 'Comportement violent de la part de Mme NAME, m'a poussé violemment sur le fauteuil [...]'. At the bottom, there are icons for document, search, and folder.

Figure 1. DEI de l'ensemble positif de la dimension Violence



The screenshot shows a form titled 'Description de l'événement'. At the top, there is a text input field containing 'Le 25/05 au diner nous a été s...'. Below this, there are several filters: 'Date' (25 mai 2013), 'Heure' (18h-20h), 'Etablissement' (empty), 'Pôle/Service' (TOUTES LES ACTIVITES), 'Processus / Activité' (BFS50), 'Causes' (empty), 'Risque' (BFS50-R6-GRI-G2), 'Gravité' (empty), and 'Matière' (empty). Below the filters, there is a 'Description de l'événement' section with a text area containing the following text: 'Le 25/05 au diner nous a été servie une omelette. Au cours des 15h suivant le repas, j'ai été prise de violentes douleurs abdominales associées à des diarrhées importantes, nausées et céphalées. Les mêmes symptômes étaient apparus durant avril après avoir mangé des œufs mayonnaise'. At the bottom right, there is a 'Fermer' button.

Figure 2. DEI de l'ensemble négatif de la dimension Violence

L'expert est encouragé à regrouper ce type de DEI dans l'« ensemble négatif » qui, avec l'ensemble positif, constitue le corpus d'apprentissage global. Cet ensemble négatif peut regrouper deux types de documents : ceux n'évoquant pas la dimension recherchée et sans lien particulier avec elle, mais aussi et surtout ceux qui, de par les termes utilisés, pourraient laisser penser à la présence de la dimension alors que ce n'est pas le cas, comme la DEI de la Figure 2. Pour reprendre l'exemple de l'expression « coup de tête » citée plus haut : il sera particulièrement intéressant de considérer dans l'ensemble négatif les DEI dans lesquelles ce groupe de mots signale l'impulsivité, et d'affilier à l'ensemble positif celles où il réfère à un acte violent.

Une fois les ensembles créés, l'utilisateur décide lui-même de lancer l'entraînement pour modéliser la dimension recherchée (Figure 3). L'entraînement se fait par classification supervisée probabiliste à deux classes (une pour l'ensemble positif, une pour l'ensemble négatif). On utilise une régression linéaire binaire¹⁰ qui fournira, pour chaque DEI, la probabilité d'appartenir à la classe positive. Les traits utilisés pour l'entraînement notamment en compte les stems et les entités nommées reconnues lors des étapes précédentes ainsi que leur agencement dans le document.

¹⁰ LibLinear (Fan et al, 2008).

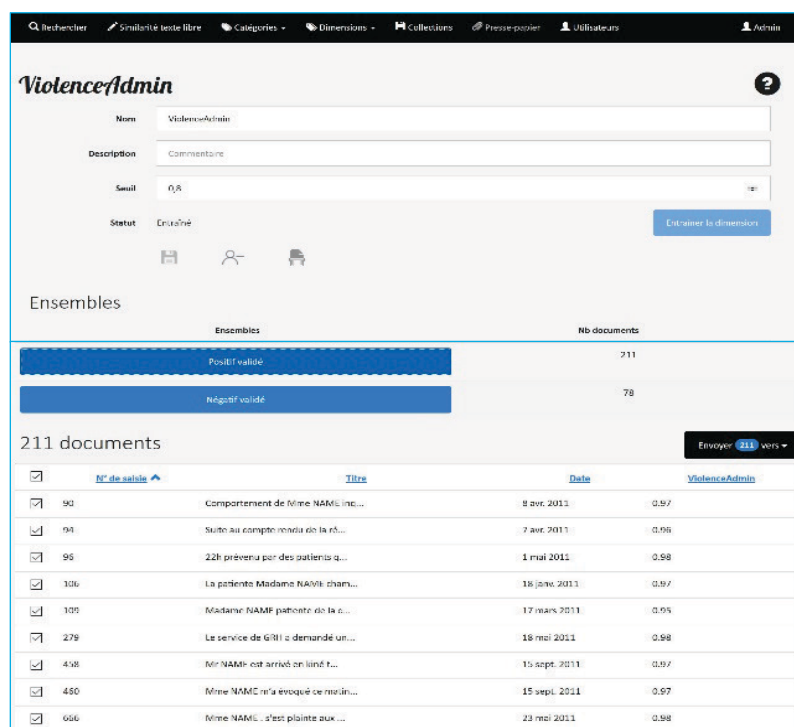


Figure 3. Interface de gestion de la dimension Violence

L'analyse automatique comparative de ces ensembles va permettre d'identifier les traits discriminants pour la dimension de celui qui ne l'est pas afin de prédire l'appartenance à l'un ou à l'autre des ensembles des documents qui seront analysés ensuite. Ainsi, à l'issue de l'apprentissage, toutes les DEI sont passées au crible du modèle qui leur attribue un poids entre 0 et 1 selon leur contenu textuel afin de signaler la plus ou moins forte présence de la dimension dans les textes.

Lorsqu'un poids a été attribué à toutes les DEI de la base de données, la métadonnée correspondant à la dimension apparaît dans l'interface de visualisation de chaque document mais également dans les critères disponibles pour la recherche, le tri et le filtrage des résultats. Par

conséquent, les experts pourront créer une requête en spécifiant un critère tel « Violence est supérieure à 0,9 » pour retrouver toutes les DEI pour lesquelles l'algorithme suggère une forte présence de la violence (cf. Figure 4). Notons avant de présenter les résultats sur la dimension « Violence » que l'apprentissage fait ici est un « apprentissage actif » (*active learning*) : l'utilisateur voit immédiatement les résultats, peut compléter les ensembles et relancer l'apprentissage autant qu'il le souhaite jusqu'à être satisfait des poids obtenus sur les documents ; cette question de la satisfaction pose immédiatement celle des résultats et de leur évaluation.

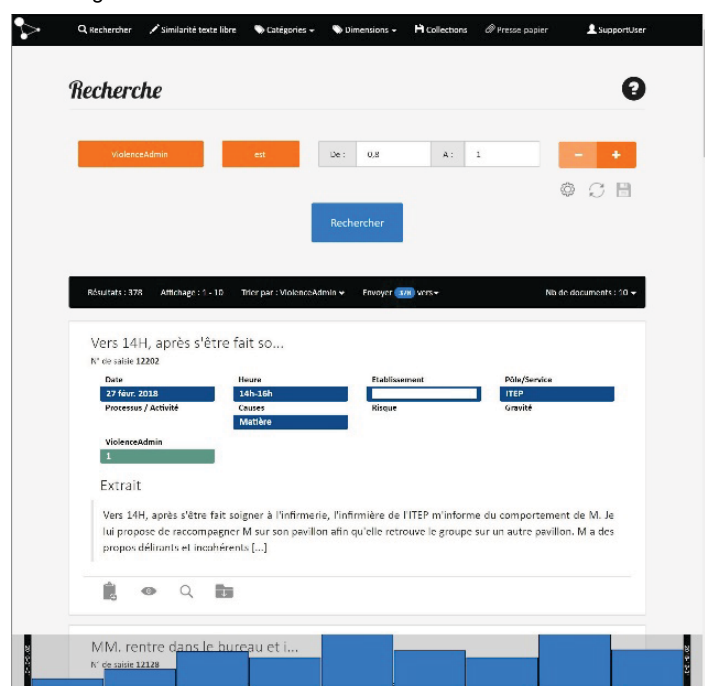


Figure 4. Recherche sur la dimension Violence et visualisation dans les documents

Résultats et évaluation

La question qui se pose dès la création d'une dimension est celle de son évaluation : il est crucial de connaître la fiabilité du poids attribué aux documents afin de savoir le degré de confiance que l'on peut avoir dans cette métrique. Cette évaluation est loin d'être triviale puisque la dimension a justement pour objectif de faire émerger des documents qui ne sont pas facilement repérables. Pour présenter la démarche de l'évaluation, nous partons du cas concret de la dimension Violence créée par l'UGECAM.

Le module de dimension a été mis en service fin 2017 et la dimension Violence créée par les experts UGECAM dans l'idée de surveiller les événements liés de près ou de loin à la maltraitance. Nous rendons ici compte des résultats obtenus au cours des 4 premiers mois d'utilisation.

Cette dimension Violence a été initiée à partir d'un ensemble positif contenant 25 DEI et d'un ensemble négatif de 22 DEI. Une fois l'apprentissage réalisé, un poids a été automatiquement attribué à l'intégralité des déclarations de la base (soit 10 846 DEI fin 2017), et la métadonnée « Violence » a été créée et est désormais visible et interrogeable via l'interface de recherche intégrée à PLUS.

En cherchant alors les DEI dont « la dimension Violence est supérieure à 0,75 », on trouve 93 DEI dans lesquelles se trouvent logiquement les 25 DEI de l'ensemble positif d'apprentissage, par conséquent, en les retranchant des résultats, on observe que 68 nouvelles DEI ont été identifiées comme porteuses de la dimension Violence.

Il est intéressant de noter que parmi elles, 55 DEI ne contiennent ni « violence » (et ses dérivés comme « violemment » ou « violent »), ni « agresser » (et ses dérivés), termes-clés qui avaient servi à constituer les ensembles du corpus d'apprentissage. On trouve en revanche dans les résultats les exemples suivants :

« [Le patient] a été très virulent et incorrect avec Madame ASH du service Restauration. »

ou

« Je suis venue l'aider pour contenir physiquement NOM qui était hors d'elle et ne voulait pas revenir à L'ETAB. »

On notera la difficulté de repérer ce type de données via une recherche classique et l'intérêt d'utiliser la dimension.

Toutefois, à l'analyse des résultats, on trouve des DEI non pertinentes (résultats faux positifs) comme par exemple :

« charge de travail trop importante et difficilement réalisable, usure morale et physique. »

Pour avoir une idée précise de la qualité des résultats, chacune des 68 DEI dont Violence est supérieure à 0,75 et n'appartenant pas aux ensembles positif et négatif a été évaluée par un expert de l'UGECAM ; la Table 1 rend compte de cette évaluation.

	0,75 ≤ d < 0,8	0,8 ≤ d < 0,9	0,9 ≤ d ≤ 1	Total
Pertinent	19,1 %	33,8 %	5,9 %	58,8 %
Non pertinent	23,5 %	17,7 %	0 %	41,2 %
Total (68 DEI)	42,6 %	51,5 %	5,9 %	100 %

Table 1. Évaluation des résultats de la dimension Violence ≥ 0,75 après le 1^{er} apprentissage

Il apparaît clairement que le seuil de 0,75 est trop permissif et qu'il est préférable de ne considérer que les DEI dont la dimension est égale ou supérieure à 0,8 : en effet, on préfère dans un premier temps améliorer la précision. En considérant les résultats sous l'angle de ce nouveau seuil de 0,8, on obtient la répartition décrite dans la Table 2.

	0,8 ≤ d < 0,9	0,9 ≤ d ≤ 1	Total
Pertinent	59 %	10,2 %	69,2 %
Non pertinent	30,8 %	0 %	30,8 %
Total (39 DEI)	89,8 %	10,2 %	100 %

Table 2. Évaluation des résultats de la dimension Violence ≥ 0,8 à l'issue du 1^{er} apprentissage

Même en réduisant l'intervalle pertinent, les résultats nécessitent d'être améliorés ; il est par conséquent nécessaire d'ajuster le corpus d'apprentissage.

La méthodologie d'ajustement consiste à analyser qualitativement les DEI les plus porteuses de la dimension et de compléter l'ensemble négatif avec les résultats non pertinents (comme celui cité précédemment), et inversement, à analyser les DEI les moins porteuses de la dimension en ajoutant dans l'ensemble positif celles où la violence est présente malgré tout.

Notons qu'il est très facile d'ajuster les ensembles : à la lecture d'un texte (que l'on soit en train d'analyser les résultats de la dimension ou en train de faire une recherche sur tout autre chose), l'utilisateur clique sur la dimension pour faire apparaître une interface lui permettant de préciser dans quel ensemble il souhaite intégrer le document. Il peut aussi bien choisir d'intégrer le document à un ensemble que de le changer d'ensemble ou même de le supprimer de l'ensemble auquel il appartient pour qu'il ne soit pas dans le corpus d'apprentissage. Il est également possible d'utiliser le « Presse-papier » disponible dans PLUS pour stocker plusieurs DEI qui seront toutes associées ultérieurement, d'un seul clic, à un même ensemble.

À ce jour, deux itérations ont été réalisées par l'expert de l'UGECAM en charge de la gestion de la dimension Violence afin de faire évoluer les ensembles et ainsi d'améliorer les résultats. Les deux tableaux ci-dessous rendent compte de l'évolution des résultats.

L'évaluation présentée dans la Table 3 est celle des résultats obtenus suite au deuxième apprentissage lancé après que le corpus a été complété et atteigne 156 DEI positives et 35 négatives ; là encore seules les DEI n'appartenant pas à ces corpus sont considérées.

	0,8 ≤ d < 0,9	0,9 ≤ d ≤ 1	Total
Pertinent	28,1%	52,6%	80,7%
Non pertinent	11,7%	7,6%	19,3%
Total (171 DEI)	39,8%	60,2%	100,0%

Table 3. Évaluation des résultats de la dimension Violence ≥ 0,8 après le 2^e apprentissage

On observe un gain très net dans la qualité des résultats ; néanmoins, un cinquième des résultats étant jugé non pertinent et l'ensemble négatif étant très peu fourni, l'expert a considéré qu'une nouvelle amélioration était possible. La Table 4 rend ainsi compte des résultats obtenus après un nouvel ajustement des ensembles les amenant à 211 DEI pour l'ensemble positif et 77 DEI pour le négatif ; là encore, seules les DEI n'appartenant pas à ces corpus sont considérées.

	0,8 ≤ d < 0,9	0,9 ≤ d ≤ 1	Total
Pertinent	30,2 %	61,5 %	91,7 %
Non pertinent	7,7 %	0,6 %	8,3 %
Total (169 DEI)	37,9 %	62,1 %	100 %

Table 4. Évaluation des résultats de la dimension
Violence $\geq 0,8$ après le 3^e apprentissage

La première constatation quant aux faux positifs est donc que leur proportion se réduit considérablement pour tomber à 3% dès lors que les corpus sont consolidés, notamment l'ensemble négatif. Cela s'explique par le fait que l'utilisateur explicite les contextes qui ne sont pas liés à la violence et que par conséquent, le modèle prend mieux en considération les particularités linguistiques de l'expression des actes considérés violents et celles de ceux qui ne le sont pas.

À cette première évaluation doit être ajoutée celle des faux négatifs, soit les cas de DEI pertinentes mais ayant un poids faible donc non repérées par le modèle de la dimension. Cette étape est plus délicate car la recherche moins aisée, par définition. Ainsi, en considérant les résultats issus du dernier apprentissage, on trouve 9 551 DEI dont la dimension Violence est inférieure à 0,4, soit près de 85% de l'intégralité des DEI. Étant donné ce volume, il n'est pas possible d'évaluer exhaustivement ces milliers de DEI pour s'assurer qu'aucune d'entre elles n'est porteuse de violence. Par conséquent, nous nous concentrons sur celles qui contiennent des termes très significatifs comme « *violent* », « *agression* », « *insulte* » et « *altercation* » (et leurs dérivés). La recherche renvoie uniquement 16 résultats ; parmi eux, on trouve 7 faux négatifs, soit des DEI qui sont porteuses de violence, mais aussi des vrais négatifs comme :

« Patiente en colère et angoissée. Elle nous dit ne pas avoir eu d'altercation avec d'autres patients. »

ou encore :

« lors de la visite pour l'entretien annuel un produit agressif est tombé sur le sol d'une chambre froide. Sol abîmé. »

Une fois le modèle stabilisé et les poids attribués jugés satisfaisants par l'expert en charge de la définition de la dimension, celui-ci peut publier la dimension afin que cette nouvelle métrique soit visible et interrogeable par l'ensemble des utilisateurs du portail. Notons qu'à l'UGECAM ALPC et Centre, un mode opératoire sur la création des dimensions est défini et reprend les étapes décrites ici pour permettre aux RAQ de s'assurer de la pertinence de la dimension créée et d'ajuster les ensembles positifs et négatifs si nécessaire.

Résultats corrélés à la réalité de terrain

Le résultat de la recherche des DEI porteuses de la dimension Violence ($\geq 0,8$) fournit un histogramme indiquant l'effectif par année (cf. Figure 4) grâce auquel on observe que le nombre de déclarations augmente jusqu'en 2014 (passant de 9 DEI en 2010 à 64 DEI en 2014).

Sur le terrain, le premier référentiel édité pour signaler les actes de maltraitance sur les patients date de 2008. Jusqu'en 2012, les préoccupations sur ce sujet ont été constantes puisque six versions du référentiel ont été diffusées. En 2014, des fiches réflexes ont été formalisées pour élargir le champ de la maltraitance aux actes à l'encontre du personnel. Il est intéressant de constater que ces recommandations n'ont pas eu pour effet, du moins dans un premier temps, de diminuer le nombre d'actes de maltraitance mais ont sensibilisé les personnels à la déclaration, ce qui a entraîné un nombre de constats plus important. Fin 2015, alors que l'on peut constater une diminution globale du nombre de déclarations (certainement en raison d'un essoufflement, ce constat étant généralisé à l'ensemble des DEI), une évaluation sur cette problématique de la maltraitance a été déployée donnant lieu en 2016 à la mise en place de nouveaux dispositifs de lutte contre la maltraitance et de promotion de la bientraitance (notamment la formation de référents bientraitance dans les établissements). En 2017, 70 déclarations sont porteuses de la dimension Violence et

sur les deux premiers mois de 2018, déjà 52 déclarations y sont associées. Il apparaît là encore que le fait de travailler sur le sujet améliore la sensibilisation des professionnels et augmente le nombre de déclarations, permettant ainsi d'avoir une vision plus exhaustive de la réalité de terrain. Ainsi, il est possible de mieux appréhender le phénomène et proposer des plans d'action plus adaptés pour l'endiguer.

En analysant à la lumière des causes (un champ « Cause » est associé à chaque DEI) les DEI dont la dimension Violence est significative, on observe qu'elles sont, comme on s'y attend, liées presque exclusivement à une origine humaine (plutôt que matérielle ou environnementale). Mais l'analyse transverse permise par la dimension permet de constater la répartition de cette origine : dans 88% la cause est liée aux patients et à leur entourage, les presque 12% restants ayant pour origine le personnel. La pertinence de la dimension est aussi confirmée par la répartition des DEI dans les processus (11 processus représentés sur les 12 existants à l'UGECAM) ainsi que dans les activités (36 activités sur les 112 existantes) et dans les risques (44 risques sur les 510 définis). Sans la dimension, le repérage de ces DEI n'aurait pas été possible.

Enfin, cette première analyse tend à montrer que le respect des libertés individuelles est conditionné au devoir de respect de l'autre. En effet, dans les DEI identifiées comme les plus violentes, on voit apparaître un phénomène de « réactions en chaîne » : les personnes agressées devenant à leur tour les « agresseurs », comme on le perçoit dans l'exemple suivant :

« Suite au compte rendu de la réunion du CHSCT qui a déplu à beaucoup de monde l'IDE Mr NAME m'a insulté me traitant de Bête et pas intelligente. »

Cette analyse doit servir de base aux prochains travaux engagés par l'UGECAM sur le sujet.

Conclusion

Plusieurs points émergent de cette étude et les conclusions que l'on peut tirer seront à réexaminer à la lumière de l'expérience que les experts feront de la dimension au cours de 2018. En effet, au cours de cette année, l'UGECAM souhaite relancer une évaluation de la maltraitance et la dimension va permettre de dresser un premier état des lieux sur la réalité de terrain afin de cibler les champs à évaluer.

D'un point de vue tant fonctionnel que qualitatif, on a observé que le seuil à considérer pour avoir les résultats les plus satisfaisants est celui de 0,8. Il permet de limiter radicalement le bruit tout en ayant une proportion de silence acceptable pour les experts. Il sera intéressant de confirmer ce seuil grâce à une analyse similaire sur une ou plusieurs autres dimensions. On travaillera également à ajuster les paramètres techniques de l'apprentissage (la méthode notamment) pour maximiser à la fois la qualité et la stabilité des résultats. Enfin, l'étude présentée ici sera complétée dans les mois qui viennent par une analyse fine et exhaustive de l'impact du volume des ensembles positifs et négatifs dans l'apprentissage. En effet, si on peut voir que l'ajout de documents négatifs améliore l'apprentissage, il conviendra de vérifier s'il est préférable qu'il soit totalement équivalent à l'ensemble positif ou jusqu'où sa moindre représentation est acceptable. De manière plus large, il s'agira également d'essayer d'évaluer, proportionnellement à l'intégralité de la base de données, combien de documents positifs et négatifs sont nécessaires pour obtenir des résultats satisfaisants ; l'enjeu étant de trouver le compromis acceptable entre le coût lié à la constitution des ensembles (qui demande un investissement de la part de l'expert) et la qualité des résultats, et par conséquent le bénéfice pour les experts dans leur travail d'analyse.

Remerciements

Les auteurs tiennent à remercier la direction de l'UGECAM ALPC et Centre qui soutient la démarche de réflexion ainsi que l'ensemble des responsables qualité de la structure pour leur contribution au cours de la formation au module de dimension de *PLUS* et des séances de travail dédiées.

Références

Amalberti, R., 1996, La conduite des systèmes à risque, PUF.

Amalberti, R., 2012, Piloter la sécurité : Théories et pratiques sur les compromis et les arbitrages nécessaires, Springer.

Fan, R. E., Chang, K. W., Hsieh, C. J., Wang, X. R., & Lin, C. J., 2008, « LIBLINEAR: A library for large linear classification. » Journal of machine learning research, 9(Aug), pages 1871-1874.

Gaucher, C. & Raynal, C., 2018, « Utilisation du TAL pour la consolidation de données de fiabilité issues des REX »,

Actes du Congrès λμ 21, « Maîtrise des risques et transformation numérique : opportunités et menaces », Reims, octobre 2018.

Lagarde, S., Andreys, J.-B., Raynal, C. & Hermann, E., 2015, « De la déclaration d'événement indésirable à la gestion des risques ; illustration d'une démarche », Techniques hospitalières, no 752, pages 31-35, 01/07/2015

Porter, M., 2001, « Snowball: A language for stemming algorithms. » <http://snowball.tartarus.org/texts/>

Tanguy, L., Tulechki, N., Urieli, A., Hermann, E. & Raynal, C., 2015, « Natural language processing for aviation safety reports: From classification to interactive analysis », Computers in Industry, Elsevier.

Urieli, A., 2013, Robust French syntax analysis: reconciling statistical methods and linguistic knowledge in the Talismane toolkit (Thèse de doctorat), Université de Toulouse II – Le Mirail.