

Repérer et modéliser un problème technique : application aux rapports d'incidents du domaine spatial

Mariame Maarouf

Directeur de thèse : Ludovic Tanguy

Encadrants CNES : Daniel Galarreta, Pascal Noir, Michal Kurela

Encadrant MeetSYS : Jérôme Laforcade

Journées CNES Jeunes Chercheurs 2 - 11, 12, 13 octobre 2023

1. Introduction

Traitement qualitatif de REX (Retours d'EXpérience) issues du CNES (Centre National d'Études Spatiales) avec des techniques de TAL (Traitement Automatique des Langues). Le cadre ici présent est celui du traitement automatique des FT (Fiches Techniques) qui rendent compte d'incidents rencontrés en phase d'exploitation de lanceurs d'Ariane 5.

Problématique : Comment repérer et extraire automatiquement, au sein d'un énoncé de type REX, les éléments qui composent le problème rencontré, la cause de ce problème, et les formaliser sous forme de relation fonctionnelle

Hypothèse : Ces énoncés font état de patterns langagiers, à partir desquels une grammaire d'expression d'un problème technique peut être distinguée.

Objectifs :

- retrouver des problèmes similaires déjà rencontrés (et résolus) dans une base documentaire,
- retrouver quelle(s) solution(s) a/ont été employée(s),
- trouver des solutions supplémentaires grâce à une méthode de résolution de problème (TRIZ), (Pour des raisons de confidentialité, les exemples ont été modifiés)

2. Processus de modélisation visé

Passage d'un texte non-contrôlé à deux étapes de modélisation sous forme de *vépoles*, formalisation issue de la méthode TRIZ (Ilevbare et al., 2013) (Altshuller et al., 1996). Identification dans le texte brut de la cause du problème, du composant altéré et du type d'interaction qui a eu lieu. Identification d'un équivalent abstrait de ces éléments.

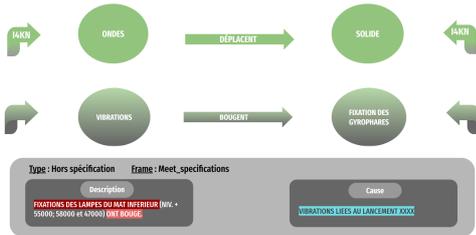


Fig. 1: Schéma du processus de modélisation

4. Spécificités des données

DEFAUT BETON SUR CARNEAUX -1) DESSUS CARNEAU XXX.JPG2) DESSUS CARNEAU NORD (1).JPG (AFFAISEMENT GENERAL)3) DESSUS CARNEAU XXX.JPG

DEFAUT DE PLANEITE SUR LE CABLE X POINTS. ON CONSTATE UN LÉGER SURPLUS (ENV. X,Xmm) SUR LA SURFACE DU CONNECTEUR (CABLE XXX COTE OXYGENE : XXXXXXXXXXX)

defaut vidange eclateur N°10 piscine XXX

- Espaces et ponctuation
- Fautes
- Minuscules
- Accents
- Acronymes



Pré-traitement / nettoyage des données avec outils traditionnels trop lourd (Bokinsky et al., 2015), (Dima et al., 2021)

6. Méthode d'annotation en deux niveaux

Annotation des données pour entraîner un modèle d'apprentissage supervisé en deux étapes.

1. indices lexicaux *déclencheurs* d'un *frame* (une situation prototypique)
2. les *participants*, éléments de l'énoncé qui participent à la situation.

Cette annotation s'appuie sur les jeux d'étiquettes définis dans la ressource *Framenet* (Ruppenhofer et al., 2010)

1. Annotation de niveau 1

- 1.1 Repérage des déclencheurs (indices lexicaux)
- 1.2 Typage des déclencheurs

1 DEGRADATION IMPORTANTE DE LA BAIE DE L'ANALYSEUR XXXXXXXX
Type : Dégradation - Usure - Saleté

2. Annotation de niveau 2

- 2.1 Repérage des participants du *frame*
- 2.2 Typage des participants

2 DEGRADATION IMPORTANTE DE LA BAIE DE L'ANALYSEUR XXXXXXXX

3. Cadre théorique : la sémantique des cadres (*frames semantics*) de (Fillmore, 1976)

Hypothèse : la compréhension de la signification d'un mot repose sur la prise en compte de leur contexte situationnel et événementiel.

Définit : des rôles ou étiquettes pour les différents éléments de l'énoncé.

L'affectation de rôles aux différents participants permet d'avoir une représentation sémantique d'un énoncé, c'est-à-dire de mettre en évidence *qui a fait quoi et comment*.

Exemple :

Jean vend une voiture en Auvergne.

Frame = vente / **Participant** = « Jean » (vendeur) / **Participant** = « voiture » (bien vendu) / **Propriété** = « Auvergne » (lieu)

5. Typologie d'expression d'un problème technique

Une typologie d'expression d'un problème technique a été construite à partir de l'étude du corpus. Cette étude a permis d'identifier neuf types d'expressions, auxquels sont associés pour chacun un ensemble marqueurs lexicaux. Ces types peuvent être mis en correspondance avec des *frames* (Fillmore, 1976).

N°	Type	Frame	Exemples de déclencheurs
1	Fuite	Fluidic_motion	fuite, fuyard, écoulement
2	Signal qui s'est déclenché	Warning	témoïn, alerte, alarme
3	Obstacle	Hindering	gêne, empêche, bloque
4	Dégradation - Usure - Saleté	Damaging	cassé, marques, corrosion
5	Élément manquant	Presence	absence, sans, manque
6	Configuration hors spécification	Meet_specifications	hors spécification, hors famille, attendu... mesuré
7	Dispositif qui ne fonctionne pas	Being_operational	HS, panne
8	Action difficile ou impossible	Difficulty	impossible, difficulté
9	État du monde		

Fig. 4: Les neuf types d'expression d'un problème technique

7. References

- Ilevbare, I. M., Probert, D., & Pahal, R. (2013). A review of triz, and its benefits and challenges in practice. *Technovation*, 33(2), 30–37. <https://doi.org/https://doi.org/10.1016/j.technovation.2012.11.003>
- Altshuller, G., Al'tov, G., & Altov, H. (1996). *And suddenly the inventor appeared: Triz, the theory of inventive problem solving*. Technical Innovation Center, Inc.
- Ruppenhofer, J., Ellsworth, M., Petruck, M., Johnson, C., & Scheffczyk, J. (2010). *Framenet ii: Extended theory and practice*.
- Fillmore, C. J. (1976). Frame semantics and the nature of language. *Annals of the New York Academy of Sciences*, 280(1), 20–32. <https://doi.org/https://doi.org/10.1111/j.1749-6632.1976.tb25467.x>
- Kurela, M., Bacqué, M., & Laurent, R. (2020). Classification automatique des faits techniques pour la conformité des lanceurs spatiaux. *Congrès Lambda Mu 22 " Les risques au cœur des transitions "* (e-congrès) - 22e Congrès de Maîtrise des Risques et de Sécurité de Fonctionnement, Institut pour la Maîtrise des Risques. <https://hal.archives-ouvertes.fr/hal-03477771>
- Dima, A., Lukens, S., Hodkiewicz, M., Sexton, T., & Brundage, M. P. (2021). Adapting natural language processing for technical text. *Applied AI Letters*, 2(3), e33. <https://doi.org/https://doi.org/10.1002/ail2.33>
- Bokinsky, H., McKenzie, A., Bayoumi, A., McCaslin, R., Patterson, A., Matthews, M. M., Schmidley, J., & Eisner, L. (2013). Application of natural language processing techniques to marine v-22 maintenance data for populating a cbm-oriented database. *AHS Airworthiness, CBM, and HUMS Specialists' Meeting*.