

# Extraction de tâches dans les e-mails : une approche fondée sur les rôles sémantiques

Melissa Mekaoui, Guillaume Tisserant, Mathieu Dodard, Cédric Lopez

Emvista

<https://www.emvista.com>

Cap Oméga, Rond-point Benjamin Franklin, 34960 Montpellier

[prenom.nom@emvista.com](mailto:prenom.nom@emvista.com)

**Résumé.** Il est estimé qu'en 2019 environ 1,4 milliard d'e-mails sont envoyés chaque jour en France (293 milliards dans le monde). Les e-mails augmentent sensiblement le volume des communications dans les entreprises. Il en découle une difficulté pour les salariés de consulter la totalité des messages reçus afin d'identifier les tâches à mener. À la fin des années 1990, les premiers systèmes visant à repérer les tâches dans les e-mails ont fait leur apparition. De nombreux travaux ont été réalisés à ce sujet (toujours d'actualité) fondés sur des méthodes d'apprentissage, des méthodes symboliques, et des méthodes hybrides. Deux approches sont communément adoptées : 1) classification d'actes de langages au niveau du message dans son intégralité ou de ses segments textuels, 2) extraction d'informations fondée sur des patrons morphosyntaxiques. Nous proposons et expérimentons une nouvelle approche fondée sur les rôles sémantiques afin de repérer les tâches dans les e-mails et de les structurer. L'évaluation de notre système sur des e-mails professionnels montre la pertinence notre proposition.

## 1 Introduction

Le substantif "tâche" est défini par le CNRTL<sup>1</sup> ainsi : "Travail défini et limité, imposé par autrui ou par soi-même, à exécuter dans certaines conditions". En contexte industriel, de nombreuses tâches sont décrites textuellement à travers des outils de communication interne tels que Slack<sup>2</sup> ou Teams<sup>3</sup>. Ces outils n'ayant évidemment pas pour objectif de véhiculer uniquement les tâches, celles-ci se trouvent noyées dans une multitude d'informations que l'utilisateur doit être en mesure de traiter.

À la fin des années 1990, les premiers systèmes visant à repérer les tâches dans les e-mails ont fait leur apparition. De nombreux travaux ont été réalisés à ce sujet, fondés sur des méthodes d'apprentissage (généralement un SVM), des méthodes symboliques (utilisant des informations linguistiques, généralement au niveau syntaxique), ou des méthodes hybrides (apprentissage + traits linguistiques).

---

1. <https://www.cnrtl.fr/definition/t%C3%A0che>

2. <https://slack.com/>

3. <https://teams.microsoft.com/>

L'originalité de notre travail consiste à repérer et à structurer les tâches par le biais d'un système d'annotation en rôles sémantiques (*Semantic Role Labeling* ou *SRL*). Les systèmes de SRL ont pour objectif d'identifier les rôles sémantiques des arguments attachés à un prédicat ; rapporté à notre contexte, il s'agit d'identifier les sens (c'est-à-dire les rôles) des éléments textuels qui constituent l'énoncé d'une tâche. Le système doit donc être en mesure de repérer qui doit faire quoi, quand, pour qui, avec quoi, *etc.* Par exemple, dans la phrase *Luc, peux-tu envoyer la facture à Léa avant demain ?*, le système doit identifier que "Luc" (rôle : Agent) devrait réaliser l'action "envoyer", que l'objet à envoyer est "la facture" (rôle : Theme) et que "Léa" (rôle : Recipient) est la destinataire de la facture.

Dans cet article, nous mettons en exergue les limites des travaux antérieurs (*cf.* section 2) puis nous définissons l'objectif du système (*cf.* section 3.1) et montrons comment l'utilisation des rôles sémantiques peut être pertinente dans ce contexte (*cf.* section 3.2). Ensuite, nous décrivons le système qui adopte une approche symbolique (*cf.* section 4) et nous l'évaluons (*cf.* section 5) avant d'en tirer quelques conclusions et perspectives (*cf.* section 6).

## 2 Travaux antérieurs

Les e-mails sont une source d'informations et de connaissances de grand intérêt pour les professionnels. Leurs contenus textuels peuvent être automatiquement exploités pour générer un historique des connaissances qu'ils véhiculent (Matta et al., 2014), pour récupérer des éléments de résolutions de problèmes redondants (Francois et al., 2016), ou encore pour identifier les tâches afin d'aider l'utilisateur dans la gestion de son temps (Khosravi et Wilks, 1999).

L'identification de tâches dans les e-mails a fait l'objet de nombreuses recherches depuis la fin des années 1990 (Khosravi et Wilks, 1999), notamment pour l'analyse de l'anglais qui a très tôt bénéficié du corpus Enron (250 000 e-mails envoyés ou reçus par 87 000 employés de la société Enron) (Klimt et Yang, 2004).

Les travaux en anglais considèrent généralement l'identification des tâches comme un problème de classification d'actes de langages : approches statistiques basées sur les n-grams (Carvalho et Cohen, 2006), algorithmes d'apprentissage supervisés (Khoo et Albrecht, 2006) (Lampert et al., 2010), non supervisés (Joty et al., 2011) (Lampert et al., 2014), semi-supervisés (Jeong et al., 2009), ou à base de règles (Scerri et al., 2010). Des traits de surfaces et des traits linguistiques (d'ordre syntaxique et lexical) sont couramment utilisés (voir notamment (Corston-Oliver et al., 2004) (Qadir et Riloff, 2011) (Lampert et al., 2014)).

Lampert et al. (2014) indiquent à juste titre que leur analyse se situe au niveau de la phrase, bien que l'énoncé d'une tâche puisse être distribuée au sein de plusieurs phrases (voire de plusieurs e-mails, de plusieurs sources de données, ou absent de tout contexte). Pour aller au-delà des frontières imposées par les phrases, Kalia et al. (2013) résolvent les coréférences pronominales<sup>4</sup> à partir d'un analyseur syntaxique en dépendances qui repère les sujets, verbes d'actions et objets.

Les systèmes cités précédemment recherchent les tâches à deux niveaux (à l'exception de Lampert et al. (2014) qui effectuent également une analyse au niveau du paragraphe) :

- au niveau du message, par exemple Corston-Oliver et al. (2004), Jeong et al. (2009) et Lampert et al. (2014)

---

4. Dans le texte "Demande à Luc d'étudier la proposition. Il doit me faire un retour avant jeudi.", "il" coréférence avec "Luc".

- 1 I'm curious - please let me know what you think about this option.  
Request (length = 47, start = 0)
- 2 I'm curious - please let me know what you think about this option.  
Request (length = 45, start = 21)

FIG. 1 – Exemples de segments textuels repérés par un système de reconnaissance de tâches dans les e-mails d'après Lampert *et al.* (2014)

Arnaud,  
Peux-tu me dire si tu prévoies de présenter des diapos sur le projet ? Si tel est le cas, peux-tu me les envoyer avant pour que j'ajuste les miennes afin d'éviter les redondances.  
Merci  
Véronique

FIG. 2 – Exemples de segments textuels repérés par un système de reconnaissance de tâches dans les e-mails d'après Kalitvianski (2018)

— au niveau de la phrase, par exemple Qadir et Riloff (2011), Jeong *et al.* (2009) et Lampert *et al.* (2014)

Dans les deux cas, il est supposé que toutes les informations nécessaires à la réalisation de la tâche sont incluses dans le segment textuel considéré (*cf.* Fig. 1, Fig. 2). Les systèmes cités retournent donc un "bloc textuel" qui, sans analyse supplémentaire, impose une limite quant à leur exploitation automatique et particulièrement concernant la gestion des tâches.

Une analyse plus fine a été proposée par Kalia *et al.* (2013) qui utilisent un analyseur syntaxique en dépendances pour repérer les sujets, verbes d'actions et objets qui doivent respecter quelques contraintes (par exemple, le sujet doit être une personne ou une organisation<sup>5</sup>). Une limite est que les verbes d'actions sont repérés *via* une liste de verbes prédéfinie sans tenir compte du contexte; l'étude cite l'exemple "I can imagine that your family reunions are just a hoot!" ("Je peux imaginer que vos réunions de famille ne sont que du plaisir!") et indique qu'il ne s'agit pas d'une tâche car le verbe "imagine" n'est pas un verbe d'action. Pourtant, dans un autre contexte, par exemple "À toi d'imaginer un nouveau slogan pour ce produit.", "imaginer" est bien une tâche. Repérer les tâches à partir d'une liste figée de verbes d'action impose *de facto* une limite importante aux systèmes.

Concernant l'analyse du français, l'extraction de tâches dans les e-mails a suscité peu de publications scientifiques, ce qui est probablement dû au nombre limité de corpus existant pour cette langue. Dans sa thèse de doctorat, Kalitvianski (2018) indique que deux corpus d'e-mails existent pour le français, mais qu'il s'agit d'e-mails non professionnels dans le premier cas<sup>6</sup> et dans un domaine spécifique dans le second cas (Chanier *et al.*, 2009). À partir d'un corpus (non publié) d'e-mails professionnels rédigés en français, Kalitvianski (2018) a proposé un système de classification supervisé des types d'énoncé ("action", "fournir document", "travailler sur document", "demande d'informations", "invitations", *etc.*) et d'attributs ("émetteur", "at-

5. Un système de reconnaissance d'entités nommées ainsi qu'un système de résolution de coréférences pronominales sont utilisés

6. Corpus Ubuntu [www.cs.cmu.edu/~pbennett/action-item-dataset.tgz](http://www.cs.cmu.edu/~pbennett/action-item-dataset.tgz)

## Extraction de tâches dans les e-mails



FIG. 3 – Exemple d'un énoncé de tâche annoté en rôles sémantiques en vue d'être structurée

tributaire", "bénéficiaire", "justification", "instruction", "condition", *etc.*) en utilisant des traits linguistiques : des indicateurs lexicaux et morphosyntaxiques ("merci de", "svp", "prière de", *etc.*) ainsi que des indicateurs tels que le nombre de destinataires ou encore la présence de pièces jointes. Pour le français, mentionnons également les travaux de Patel et al. (2019) qui s'appuient sur l'élaboration de patrons linguistiques pour détecter les tâches dans des transcriptions de réunions. Les trois patrons définis exploitent le sujet (qui doit être un "pronom personnel ou une entité nommée"), le verbe et l'objet tout en s'appuyant sur un "vocabulaire des tâches" prédéfini. Ainsi, une tâche dont le sujet est implicite, par exemple "Veuillez organiser le séminaire le mois prochain !" ne semble pas être retenue. Une limite à cette approche est que la tâche est partiellement structurée ; par exemple, le lieu d'exécution de la tâche, la manière dont elle doit être réalisée, son bénéficiaire, sa cause et sa conséquence, entre autres, ne sont pas repérées.

Finalement, les travaux antérieurs font face à de nombreux inconvénients :

- les verbes décrivant une tâche sont repérés à partir d'une liste de verbes d'action, sans tenir compte du contexte
- les analyses sont effectuées au niveau de la phrase (sauf Kalia et al. (2013) qui traite les coréférences)
- les analyses sont effectuées en domaine fermé (comme Lampert et al. (2014) nous souhaitons une analyse en domaine ouvert)
- les analyses consistent en une classification de segments textuels : les tâches ne sont pas structurées.
- du point précédent découle le fait que peu d'approches sont en mesure de distinguer plusieurs tâches au sein d'une phrase (comme le proposent Jlalaty et al. (2018)).

Pour outrepasser ces limites, nous définissons le problème de l'extraction des tâches du point de vue de l'annotation en rôles sémantiques (*Semantic Role Labeling* ou *Shallow Semantic Parsing*) telle que définie en Traitement Automatique du Langage Naturel : il s'agit d'indiquer le rôle que chaque entité joue par rapport au prédicat, dans le contexte donné (*cf.* Fig. 3).

### 3 Définitions et rôles sémantiques

Dans cette section, nous décrivons les objectifs du système et introduisons les rôles sémantiques selon ces objectifs.

### 3.1 Définition de l'objectif

Il n'existe pas de consensus concernant la typologie des tâches à adopter qui demeure dépendante du domaine d'application. Par exemple, Freed et al. (2008) classe les tâches dans les catégories *Email reply*, *Modify room date*, *Modify event data* et Scerri et al. (2010) utilisent une ontologie pour classer des *Meeting Request*, *Task Assignment*, ou encore *FileRequest*. À partir de cette observation, Lampert et al. (2008) propose le concept "action" ayant comme sous-concepts "request" et "commitment" définis ainsi :

- *request* ("demande") : phrase qui demande au destinataire de l'e-mail de programmer/réaliser l'action décrite<sup>7</sup>
- *commitment* ("engagement") : phrase qui engage l'expéditeur (ou une autre personne) à réaliser l'action décrite

Nous considérons qu'une tâche du point de vue du destinataire de l'e-mail consiste à répondre à l'e-mail ou/et à réaliser une action. Dans la suite de l'article, nous nous focalisons sur les *request* définis par Lampert et al. (2008) (hormis le fait que nous considérons le contexte au-delà de la phrase) que nous distinguons en :

- Demandes d'action : suscitent une action de la part du destinataire autre que celle de répondre au message. Par exemple "Peux-tu envoyer la facture à Luc stp?". L'action peut être négative, par exemple "Merci de ne pas envoyer la facture pour l'instant".
- Demandes d'information : ne suscitent pas d'action de la part du destinataire autre que celle de répondre au message. Par exemple "Qu'est-il arrivé à notre collègue?".

### 3.2 Du point de vue des rôles sémantiques

En linguistique, les rôles sémantiques caractérisent le rôle de chaque argument (hors explétifs) en relation avec le prédicat. Il n'existe pas de consensus sur la liste des rôles à adopter, ni même sur leur niveau de granularité. Guidés par notre objectif de structurer les demandes (*cf.* section 3.1), nous optons pour les rôles utilisés dans VerbNet (Schuler, 2005)<sup>8</sup>.

Dans le cadre des e-mails, nous considérons qu'une action nécessite un Agent qui a pour rôle de réaliser l'action. De fait, nous considérons que l'utilisateur (qu'il soit expéditeur ou destinataire) se voit affecté une action lorsque lui-même a le rôle d'Agent. Lorsque l'utilisateur se voit affecté d'autres rôles tels que Bénéficiaire (entité recevant un profit ou un dommage par l'action réalisée, par exemple "fais-le pour *Luc*"), Recipient (entité qui reçoit une chose par la réalisation de l'action, par exemple "j'ai reçu le courrier") ou Patient (entité qui subit l'action, par exemple "elle habille *Luc*"), nous considérons que l'utilisateur est impliqué dans l'action mais n'en est pas l'acteur principal : il s'agit donc dans ce cas d'une information. D'autres rôles, tels que Time (contexte temporel d'une action), Location (contexte spatial d'une action), Purpose (but de l'action), Manner (manière dont l'action est effectuée), Instrument (entité servant involontairement à l'accomplissement d'une action), Cause (cause de l'action), ou encore Consequence (conséquence de l'action) sont pertinents pour décrire une action dans le cadre des e-mails.

7. "we define a request as an utterance from the email sender that places an obligation on a recipient to schedule an action (*e.g.*, add to a calendar or task list), perform an action, or respond." Lampert et al. (2010)

8. Les rôles proto-agent et proto-patient de Dowty (1991) paraissent trop génériques au regard de l'application visée, alors que les rôles de Fillmore (1977) sont trop précis.

## 4 Traitement automatique

Le système mis en place a pour objectif de repérer et de structurer les demandes d'actions et les demandes d'informations. Le système que nous expérimentons adopte une approche symbolique. Il est composé de deux modules : 1) un module utilisant une ressource lexicosémantique, 2) un module sans ressource lexicosémantique.

### 4.1 Module avec ressource

VerbNet est une ressource en anglais où les verbes sont groupés en différentes classes (Schuler, 2005) (exemples de classes : *transfer\_mesg-37.1.1*, *acquiesce-95*, *give-13.1*). Les verbes d'une classe donnée partagent les patrons syntaxiques auxquels sont associés les rôles sémantiques. Ainsi, l'intérêt principal de VerbNet pour l'annotation automatique en rôles sémantiques est le couplage syntaxe-sémantique. Récemment, Danlos et al. (2015) ont proposé une version française de VerbNet, nommée VerbeNet, ce qui vient combler un manque depuis longtemps remarqué (Pado et Pitel, 2007). À notre connaissance, un seul système a expérimenté l'utilisation de cette nouvelle ressource pour l'annotation en rôles sémantiques pour le français (Pradet, 2015).

Notre premier module utilise VerbeNet. Il suit de près le système de Pradet (2015), une implémentation améliorée du système décrit par Swier et Stevenson (2004). Il se décline en trois phases principales :

1. Identification des prédicats. Par défaut, nous considérons comme prédicat tous les verbes détectés par un analyseur morphosyntaxique. Les locutions verbales sont repérées en utilisant la ressource JeuxDeMots (Lafourcade, 2007).
2. Identification des arguments. Plutôt que d'utiliser la position des arguments dans la phrase comme Pradet (2015), nous utilisons les résultats d'un analyseur syntaxique en dépendances<sup>9</sup>. L'utilisation des dépendances permet de repérer les arguments quel que soit leur position (gauche ou droite) et leur distance par rapport au prédicat. Les arguments identifiés permettent de générer le cadre syntaxique de chaque prédicat. Par exemple, dans *tu devrais envoyer la facture* et *tu devrais l'envoyer*, "la facture", situé après le prédicat, et "l'" situé avant le prédicat génèrent le cadre unique NP V NP<sup>10</sup> puisque "la facture" et "l'" sont rattachés au verbe en tant qu'objet. Les cadres générés automatiquement sont comparés aux cadres syntaxiques fournies par VerbeNet.
3. Correspondance des cadres et classification des arguments. Les prédicats identifiés puis lemmatisés permettent de sélectionner zéro, une ou plusieurs classe(s) VerbeNet. Chaque classe fournit un ensemble de cadres syntaxiques. Dans VerbeNet, chaque cadre syntaxique est associé à un cadre sémantique et fournit une correspondance syntagme-rôle. Lorsqu'un cadre syntaxique VerbeNet correspond à celui de la phrase analysée<sup>11</sup>, le cadre sémantique associé est sélectionné. Plusieurs cadres sémantiques peuvent être sélectionnés.

---

9. Cela implique que le système doit construire le syntagme associé à chaque *token* annoté en rôle sémantique. Cette construction est réalisée en utilisant les relations de dépendances.

10. *Noun Phrase* : syntagme nominal

11. Cette correspondance s'effectue en terme d'identité et d'inclusion. Lorsqu'aucune correspondance n'est validée, aucun cadre n'est retenu et donc aucun rôle n'est annoté. Nous évitons ainsi une approche probabiliste (Swier et Stevenson, 2004; Pradet, 2015) qui impacte la précision.

4. Sélection du cadre le plus pertinent. Les restrictions de sélection de VerbeNet sont utilisées (par exemple la classe *transfer\_msg-37.1.1* admet un Agent dont la restriction est "animé" ("animé")). Les restrictions imposées par VerbeNet sont comparées à celles calculées par l'analyseur linguistique. Le ou les cadres retenus sont ceux qui respectent le plus ces contraintes.

Un exemple de traitement est donné pour la phrase *tu devrais envoyer la facture* :

1. Le prédicat "envoyer" est identifié à l'issue de l'analyse morphosyntaxique
2. Le cadre syntaxique généré automatiquement à partir de la phrase est *NP V NP*
3. Les classes VerbeNet qui contiennent ce verbe sont *give-13.1* et *send.11.1*
4. Chaque classe donne un ensemble de cadres possibles dont : **give-13.1** : *NP.Agent V NP.Theme {à} NP.Recipient, NP.Theme V {à} NP.Recipient*; **send.11.1** : *NP.Agent V NP.Theme {Loc-Dest} NP.Destination, NP.Agent V NP.Theme {de/depuis} NP.Initial\_Location {Loc-Dest} NP.Destination*. L'unique correspondance exacte avec le cadre généré automatiquement est *NP.Agent V NP.Theme* ce qui attribue le rôle Agent à "tu" et le rôle Theme à "facture".

Notons que le système gère la voix passive (VerbeNet encode tous les cadres à la voix active).

VerbeNet a une couverture faible<sup>12</sup> ce qui a un impact immédiat sur la couverture du système. Pour augmenter la couverture, d'une part, nous avons manuellement enrichi la ressource via une interface dédiée, d'autre part nous avons développé un second module qui n'utilise aucune ressource de type lexicosémantique.

## 4.2 Module sans ressource

Le système sans ressource est également fondé sur la sortie d'un analyseur syntaxique en dépendances et tient compte des résultats du premier module : ce module se focalise sur le traitement des rôles n'ayant pas été découverts dans le module précédent. Il a pour objectif d'augmenter la couverture au risque que la précision des résultats soit affectée. Un ensemble de règles a été développé manuellement à partir d'une analyse en corpus.

Le système est constitué de deux étapes principales :

1. Génération de candidats pour chaque rôle. Chaque règle permet d'alimenter une liste de syntagmes candidats pour chaque rôle. Deux types de règles ont été développés : 1) règles basées sur les dépendances syntaxiques 2) règles basées sur des patrons morphosyntaxiques. L'intérêt est de générer de nombreux candidats de façon à assurer une bonne couverture alors que la précision serait apportée par les restrictions dans l'étape suivante.
2. Sélection du candidat. Cette étape consiste à sélectionner un unique candidat parmi ceux générés à l'étape précédente. Un candidat est sélectionné si la sémantique qui lui a été assignée par l'analyseur linguistique correspond aux restrictions de sélection les plus probables pour chaque rôle selon VerbeNet (par exemple, un agent est très probablement une "personne" ou une "organisation" alors qu'un thème a plus rarement

12. d'après Pradet (2015) mais les résultats sont prometteurs. La ressource demeure en cours de construction (<http://verbenet.inria.fr/>).

## Extraction de tâches dans les e-mails

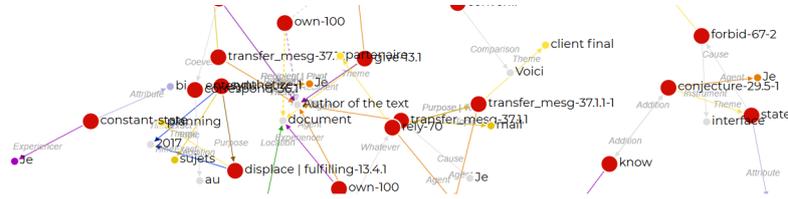


FIG. 4 – Représentation graphique des contenus textuels d'e-mails structurés par le système ; les nœuds principaux représentent les prédicats, les nœuds secondaires représentent les arguments, les relations représentent les rôles sémantiques des arguments vis-à-vis des prédicats.

ces restrictions). Dans le cas où plusieurs candidats sont pertinents, le système n'est pas en mesure de les départager et retourne plusieurs solutions.

### 4.3 Filtres

Les deux modules décrits précédemment fournissent la totalité du contenu textuel structuré en fonction des prédicats et des rôles sémantiques (*cf.* Fig. 4). Nous mettons en place des filtres qui permettent de conserver uniquement les informations souhaitées. Par exemple, nous définissons ci-après un filtre concernant les demandes d'actions telles que définies en section 3.1 :

- au niveau morphosyntaxique. Le prédicat doit être à l'impératif, ou introduit par un semi-auxiliaire (dans ce cas le prédicat est à l'infinitif), ou appartenir à une proposition subordonnée complétive conjonctive.
- au niveau des rôles. Le prédicat nécessite un Agent qui doit être le destinataire de l'e-mail.
- au niveau des classes sémantiques. Nous adoptons la proposition de Scerri et al. (2010) qui consiste à séparer les suggestions d'actions des demandes d'actions. Le prédicat ne doit donc pas être dépendant d'un autre prédicat faisant partie des classes intégrant des verbes tels que "penser", "présumer", "espérer", "croire", *etc.* (classe "conjecture-29.5-1"). Par exemple, aucune demande d'action n'est retenue dans "je pense que tu devrais gérer ce projet".

D'autres filtres peuvent être créés pour conserver d'autres types relatifs aux actions (actions suggérées, actions réalisées, actions où le destinataire est impliqué mais sans avoir à réaliser l'action, *etc.*).

### 4.4 Traitement des demandes d'information

Le système considère par défaut qu'une phrase interrogative constitue une demande. Les demandes d'informations sont définies comme étant les demandes qui ne sont pas des demandes d'actions. Ainsi, "Peux-tu envoyer le courrier stp?" est considérée comme une demande d'action par le système (les contraintes du filtre défini précédemment sont satisfaites), alors que "As-tu envoyé la facture?" est traité comme une demande d'information (il ne s'agit pas d'une demande d'action car le prédicat est au passé).

## 5 Évaluation

Dans le cadre de cette évaluation, nous avons utilisé un analyseur syntaxique fondé sur DRAGNN (Kong et al., 2017), un parser en dépendances construit à base de *deep learning*. Ce parser a été entraîné sur Sequoia (Candito et Seddah, 2012) complété d’e-mails annotés manuellement. Une surcouche symbolique permet de corriger certaines erreurs du parser. La reconnaissance des entités nommées est réalisée par le système décrit par Lopez et al. (2019). Les modules 1 et 2 (*cf.* section 4) s’appliquent consécutivement en prenant en entrée la sortie du parser enrichie par les entités nommées.

Compte tenu qu’il n’existe pas de corpus annoté en tâches disponible pour le français (*cf.* section 2), nous en avons construit un que nous avons segmenté en deux parties :

- Le jeu de développement contient 1000 e-mails. Ce jeu a été utilisé pour développer le système.
- Le jeu de test contient 200 e-mails, hors spams, invitations automatiques, newsletters, et autres e-mails envoyés automatiquement. L’e-mail le plus long contient 342 tokens, le plus court en contient 3. Ce jeu a été annoté manuellement en rôles sémantiques, selon les rôles de VerbNet. Le jeu de test a été anonymisé afin de la rendre public<sup>13</sup>.

Le corpus de test contient 87 e-mails contenant au moins une demande. Il contient 65 demandes d’information et 74 demandes d’action. Concernant les demandes d’action, 12 rôles apparaissent dans ce corpus : 74 Agent, 48 Theme/Patient, 35 Recipient, 9 Topic, 9 Purpose, 7 Cause, 5 Time, 5 Product, 4 Manner, 3 Destination, 2 Instrument, 2 Beneficiary, soient 203 annotations.

Le système a été évalué avec les mesures classiques de Précision, de Rappel et de F-mesure à trois niveaux de granularité :

1. au niveau des messages : il s’agit d’évaluer la capacité du système à détecter les e-mails contenant des demandes
2. au niveau des demandes : il s’agit d’évaluer la capacité du système à distinguer les demandes d’information des demandes d’action
3. au niveau des rôles : il s’agit d’évaluer la capacité du système à identifier les rôles associés à chaque prédicat

Niveaux	Sans correction			Avec correction		
	P	R	F	P	R	F
Messages	0,94	0,88	0,91	0,94	0,94	<b>0,94</b>
Demandes d’informations	0,98	0,89	0,93	1	0,89	<b>0,94</b>
Demandes d’actions	0,84	0,82	0,83	0,85	0,92	<b>0,88</b>
Roles	0,93	0,76	0,84	0,93	0,85	<b>0,89</b>

TAB. 1 – Résultats obtenus sur les e-mails pour chaque niveau d’analyse

Chaque niveau a été expérimenté sur le jeu de test sans correction orthographique puis avec correction orthographique réalisée manuellement pour en mesurer l’impact réel. Le tableau 1 présente les résultats pour chacun des niveaux, avec et sans correction.

13. <https://www.emvista.com/publications/>

**Au niveau des messages.** À titre d'indication, nous reportons ici les résultats obtenus par des systèmes visant également à détecter les e-mails<sup>14</sup> contenant des demandes. Le système de Carvalho et Cohen (2006) (basé sur un SVM) obtient un taux d'erreur de 17%. Le système de Lampert et al. (2014), également basé sur un classifieur SVM obtient une précision de 0,85 et un rappel de 0,84 (F-score : 0,84). Concernant les e-mails français, le système de Kalitvianski (2018) (SVM) obtient une précision de 0,84 et un rappel de 0,59 (F-score : 0,69). Notre système obtient une précision de 0,94 et un rappel de 0,88 (F-score : 0,91). L'impact de la correction orthographique est significatif : le rappel du système augmente de 0,06 (F-score : 0,94).

**Au niveau des demandes.** Le système détecte les demandes d'action avec une précision de 0,84 et un rappel de 0,82 (F-score : 0,83). Les demandes d'informations sont reconnues avec une précision de 0,98 et un rappel de 0,89 (F-score : 0,93). Les demandes d'informations et d'actions sont mieux distinguées après correction orthographique : le système obtient une précision parfaite pour les demandes d'informations avec un rappel de 0,89 (F-score : 0,94). Concernant les demandes d'actions, le système obtient une précision de 0,85 et un rappel significativement meilleur avec la correction : 0,92 (F-score : 0,88).

**Au niveau des rôles.** Le système annote les rôles avec une précision de 0,93 et un rappel de 0,76 (F-score : 0,84). Après correction orthographique, le rappel augmente significativement (+0,09) et le F-score atteint 0,89.

## 6 Conclusions

L'extraction de tâches dans les e-mails a suscité beaucoup d'attention ces vingt dernières années (cf. section 2). Les systèmes développés s'appuient majoritairement sur des méthodes de classification au niveau du message ou de la phrase sans chercher à structurer les tâches dans leur intégralité. Dans cet article, nous avons expérimenté une nouvelle approche pour extraire et structurer les tâches dans les e-mails. L'originalité de notre travail consiste à aborder le sujet du point de vue de l'extraction d'évènements en utilisant la ressource lexicosémantique VerbNet. Les résultats obtenus indiquent que le système est très performant pour identifier les e-mails contenant des demandes (F-score entre 0,91 et 0,94). De façon plus fine, le système est également en mesure de distinguer les demandes d'action des demandes d'information avec un taux de réussite très satisfaisant (F-score autour de 0,90). Enfin, les résultats obtenus concernant les rôles sont également très satisfaisants (F-score entre 0,84 et 0,89). La perspective majeure à ce travail consiste à développer un correcteur (et normaliseur) automatique qui a un impact important sur les résultats.

## Références

Candito, M. et D. Seddah (2012). Le corpus sequoia : annotation syntaxique et exploitation pour l'adaptation d'analyseur par pont lexical (the sequoia corpus : Syntactic annotation and use for a parser lexical domain adaptation method)[in french]. In *Proceedings of the Joint Conference JEP-TALN-RECITAL 2012, volume 2 : TALN*, pp. 321–334.

---

14. (Le corpus utilisé est celui d'Enron ; e-mails anglais.)

- Carvalho, V. R. et W. W. Cohen (2006). Improving email speech acts analysis via n-gram selection. In *Proceedings of the HLT-NAACL 2006 Workshop on Analyzing Conversations in Text and Speech*, pp. 35–41. Association for Computational Linguistics.
- Chanier, T., M.-N. Lamy, C. Reffay, M.-L. Betbeder, et M. Ciekanski (2009). Letec (learning and teaching corpus) simuligne.
- Corston-Oliver, S., E. Ringger, M. Gamon, et R. Campbell (2004). Task-focused summarization of email. In *Text Summarization Branches Out*, pp. 43–50.
- Danlos, L., T. Nakamura, et Q. Pradet (2015). Traduction de VerbNet vers le français. In *Congrès ACFAS*, Rimouski, Canada, pp. 1. ACFAS.
- Dowty, D. (1991). Thematic proto-roles and argument selection. *language*, 547–619.
- Fillmore, C. J. (1977). Scenes-and-frames semantics. *Linguistic structures processing* 59, 55–88.
- Francois, R., M. Nada, et A. Hassan (2016). Ktr : an approach that supports knowledge extraction from design interactions. *IFAC-PapersOnLine* 49(12), 473–478.
- Freed, M., J. G. Carbonell, G. J. Gordon, J. Hayes, B. A. Myers, D. P. Siewiorek, S. F. Smith, A. Steinfeld, et A. Tomasic (2008). Radar : A personal assistant that learns to reduce email overload. In *AAAI*, pp. 1287–1293.
- Jeong, M., C.-Y. Lin, et G. G. Lee (2009). Semi-supervised speech act recognition in emails and forums. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing : Volume 3-Volume 3*, pp. 1250–1259. Association for Computational Linguistics.
- Jlailaty, D., D. Grigori, et K. Belhajjame (2018). Email business activities extraction and annotation. In *International Workshop on Information Search, Integration, and Personalization*, pp. 69–86. Springer.
- Joty, S. R., G. Carenini, et C.-Y. Lin (2011). Unsupervised modeling of dialog acts in asynchronous conversations. In *Twenty-Second International Joint Conference on Artificial Intelligence*.
- Kalia, A., H. R. M. Nezhad, C. Bartolini, et M. Singh (2013). Identifying business tasks and commitments from email and chat conversations. *tech. report, HP Labs*.
- Kalitivianski, R. (2018). *Traitements formels et sémantiques des échanges et des documents textuels liés à des activités collaboratives*. Ph. D. thesis, Université Grenoble Alpes.
- Khoo, Y. M. A. et D. Albrecht (2006). Experiments with sentence classification. In *Proceedings of the Australasian Language Technology Workshop 2006*, pp. 18–25.
- Khosravi, H. et Y. Wilks (1999). Routing email automatically by purpose not topic. *Natural Language Engineering* 5(3), 237–250.
- Klimt, B. et Y. Yang (2004). Introducing the enron corpus. In *CEAS*.
- Kong, L., C. Alberti, D. Andor, I. Bogatyy, et D. Weiss (2017). Dragnn : A transition-based framework for dynamically connected neural networks. *arXiv preprint arXiv :1703.04474*.
- Lafourcade, M. (2007). Making people play for lexical acquisition with the jeuxdemots prototype. In *SNLP'07 : 7th international symposium on natural language processing*, pp. 7.
- Lampert, A., R. Dale, et C. Paris (2010). Detecting emails containing requests for action.

- In *Human Language Technologies : The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 984–992. Association for Computational Linguistics.
- Lampert, A., R. Dale, C. Paris, et al. (2008). The nature of requests and commitments in email messages. In *Proceedings of the AAAI Workshop on Enhanced Messaging*, pp. 42–47.
- Lampert, A. T. et al. (2014). Making email actionable : the identification and use of obligation acts in workplace email.
- Lopez, C., M. Mekaoui, K. Aubry, J. Bort, et P. Garnier (2019). Reconnaissance d’entités nommées itérative sur une structure en dépendances syntaxiques avec l’ontologie nerd. In *Extraction et Gestion des Connaissances : Actes de la conférence EGC’2019*, Volume 79, pp. 81–92. BoD-Books on Demand.
- Matta, N., H. Atifi, et F. Rauscher (2014). Knowledge extraction from professional e-mails. In *IFIP International Workshop on Artificial Intelligence for Knowledge Management*, pp. 43–57. Springer.
- Pado, S. et G. Pitel (2007). Annotation précise du français en sémantique de rôles par projection cross-linguistique. *Proceedings of TALN-07, Toulouse, France*.
- Patel, N., M. Lannes, et C. Pradel (2019). Patrons linguistiques pour l’extraction de tâches dans des transcriptions de réunions. In *Actes de PFIA 2019*, pp. 158–166. PFIA.
- Pradet, Q. (2015). *Annotation en rôles sémantiques du français en domaine spécifique*. Ph. D. thesis, Université Paris Diderot (Paris 7).
- Qadir, A. et E. Riloff (2011). Classifying sentences as speech acts in message board posts. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 748–758. Association for Computational Linguistics.
- Scerri, S., G. Gossen, B. Davis, et S. Handschuh (2010). Classifying action items for semantic email. In *LREC*.
- Schuler, K. K. (2005). Verbnet : A broad-coverage, comprehensive verb lexicon.
- Swier, R. S. et S. Stevenson (2004). Unsupervised semantic role labelling. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*.

## Summary

In 2019, around 1.4 billion e-mails are sent every day in France (293 billion worldwide). E-mails significantly increase the volume of communications in companies. As a result, it is difficult for employees to read all messages in order to identify the tasks to be carried out. First systems to identify tasks in e-mails appeared at the end of the 1990s. Much work has been done on this topic, based on machine learning, symbolic methods, and hybrid methods. Two approaches are commonly adopted: 1) classification of language acts (at the message level or sentence level, 2) information extraction based on linguistic patterns. We propose and experiment with an approach based on event extraction (from Information Extraction) and Semantic Role Labeling to identify and structure tasks in emails. The evaluation of our system on professional e-mails shows the relevance of our proposal.