

# Traitement Automatique de la Langue et Intégration de Données pour les Réunions de Concertations Pluridisciplinaires en Oncologie

Nesrine Bannour<sup>1</sup> Aurélie Névéol<sup>1</sup> Xavier Tannier<sup>2</sup> Bastien Rance<sup>3</sup>

(1) Laboratoire Interdisciplinaire des Sciences du Numérique (LISN), rue John Von Neumann, 91400 Orsay, France

(2) Laboratoire d'Informatique Médicale et d'Ingénierie des Connaissances en e-Santé (LIMICS), 15 Rue de l'École de Médecine, 75006 Paris, France

(3) Hôpital Européen Georges Pompidou (HEGP) AP-HP, 20 Rue Leblanc, 75015 Paris, France  
nesrine.bannour@limsi.fr, aurelie.neveol@limsi.fr,  
xavier.tannier@sorbonne-universite.fr, bastien.rance@aphp.fr

## RÉSUMÉ

---

Les réunions de concertations pluridisciplinaires (RCP) en oncologie permettent aux experts des différentes spécialités de choisir les meilleures options thérapeutiques pour les patients. Les données nécessaires à ces réunions sont souvent collectées manuellement, avec un risque d'erreur lors de l'extraction et un coût important pour les professionnels de santé. Plusieurs travaux scientifiques portant sur des documents en anglais se sont intéressés à l'extraction automatique d'informations (telles que la localisation de la tumeur, les classifications histologiques, TNM, ...) dans les rapports cliniques des dossiers médicaux (Nguyen *et al.*, 2010; Savova *et al.*, 2017; Gupta *et al.*, 2019). Dans le cadre du projet ASIMOV (ASsIster la recherche en oncologie par le Machine Learning, l'intégration de données et la Visualisation), nous utiliserons le traitement automatique de la langue et l'intégration de données pour l'extraction d'informations liées au cancer dans les entrepôts de données et les textes cliniques en français, comme illustré en figure 1.

**MOTS-CLÉS** : Traitement Automatique de la Langue, Intégration de Données, Analyse Temporelle, Parcours de Soins, Oncologie.

---

## Références

- GUPTA K., THAMMASUDJARIT R. & THAKKINSTIAN A. (2019). NLP automation to read radiological reports to detect the stage of cancer among lung cancer patients. In *WNL@ACL*.
- NGUYEN A., LAWLEY M., HANSEN D., BOWMAN R., CLARKE B., DUHIG E. & COLQUIST S. (2010). Symbolic rule-based classification of lung cancer stages from free-text pathology reports. *Journal of the American Medical Informatics Association : JAMIA*, **17** 4, 440–5.
- SAVOVA G., TSEYTLIN E., FINAN S., CASTINE M., MILLER T., MEDVEDEVA O., HARRIS D., HOCHHEISER H., LIN C., CHAVAN G. & JACOBSON R. S. (2017). DeepPhe - a natural language processing system for extracting cancer phenotypes from clinical records. *Cancer research*, **77** 21, e115–e118.

Compte rendu d'hospitalisation

MOTIF D'HOSPITALISATION :  
Cure de chimiothérapie n° 2 pour  
**adénocarcinome du colon droit  
avec métastases hépatiques  
synchrones.**

SCHEMA DE TRAITEMENT :  
ANAMNESE :  
EXAMEN CLINIQUE :  
**Indice de performance: grade  
OMS 0.** Poids: 74 kg. Surf. corp.:  
1.87 m2.

**Tumeur**

Tumeur : adénocarcinome du  
colon droit  
Localisation : colon droit  
Métastases : hépatiques  
Etat clinique : grade OMS 0

Compte rendu  
d'anatomo-pathologie

Conclusion :  
**adénocarcinome lieberkühnien  
du colon droit T4N1M1** avec  
**métastases hépatiques  
synchrones**

**Tumeur**

Tumeur : adénocarcinome du  
colon droit  
Localisation : colon droit  
Type histologique : lieberkühnien  
T : T4  
N : N1  
M : M1

**Tumeur**

Tumeur : adénocarcinome du colon  
droit  
Localisation : colon droit  
Type histologique : lieberkühnien  
Métastases : hépatiques  
Etat clinique : grade OMS 0  
T : T4  
N : N1  
M : M1

FIGURE 1 – Extraction d'informations cancer d'un dossier patient en français