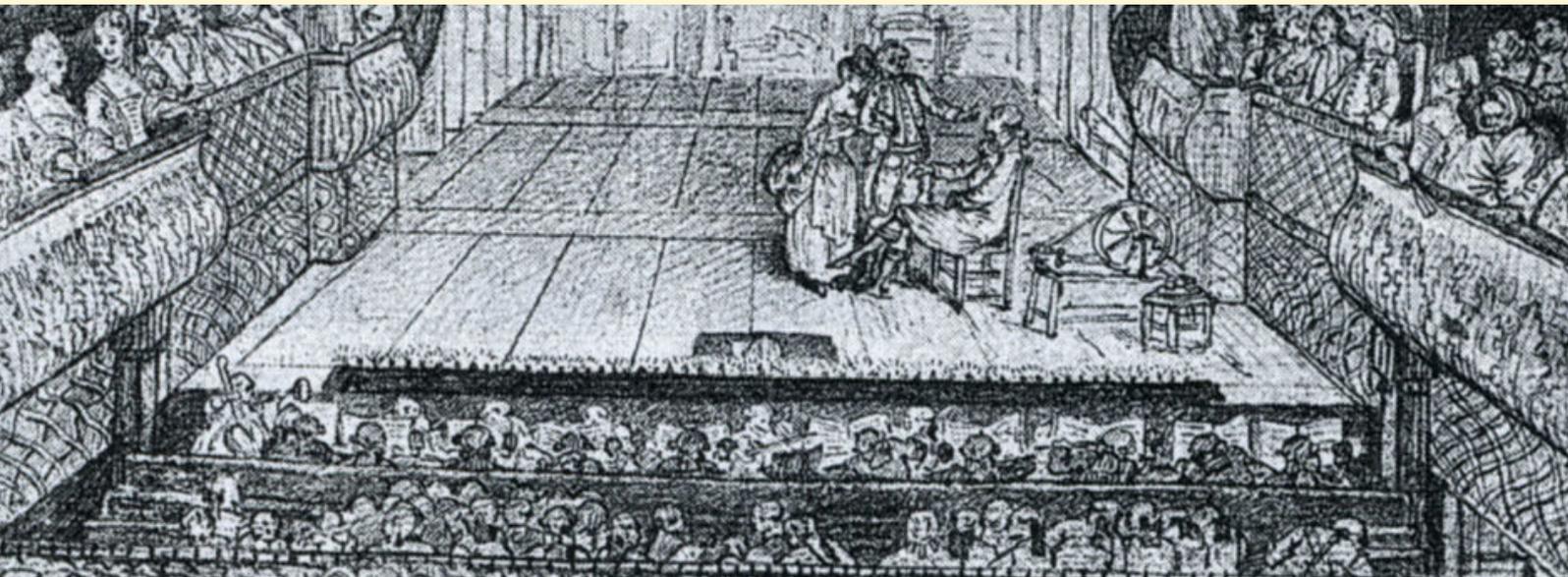


RECITAL

Des nouvelles de la Comédie-Italienne pour les crowdsourcers



JOYEUX NOËL !

Par Benjamin Hervy

Déjà le deuxième numéro pour cette lettre d'information mensuelle. Vous retrouverez au pied du sapin certaines de vos rubriques favorites, avec le déjà célèbre défi du mois !

Et comme RECITAL est avant tout le fruit d'un projet de recherche (CIRESFI qui signifie Contrainte et Intégration : pour une réévaluation des spectacles forains et italiens sous l'Ancien Régime), nous consacrons ce mois-ci la rubrique Sciences&Techno à la présentation des travaux d'Adeline Granet, doctorante en reconnaissance d'écriture automatique. Vous vous apercevrez que les documents manuscrits anciens tels que les registres comptables donnent du fil à retordre aux techniques les plus avancées des sciences du numérique !

À bientôt et très bonnes fêtes de fin d'année !

DANS CE NUMERO

LES CHIFFRES DU MOIS

SCIENCES & TECHNO

PROCHAIN DÉFI



LES CHIFFRES DU MOIS

MEILLEUR CONTRIBUTEUR

Nombre de tâches réalisées par la personne la plus active entre le 14/11 et le 19/12, bravo !

751

5827

NOUVELLES INFORMATIONS

Incluant les "marques" créées et les transcriptions réalisées.

NOMBRE DE TÂCHES DÉJÀ RÉALISÉES

Depuis la mise en ligne de la plateforme Recital. Keep going !

38226

109



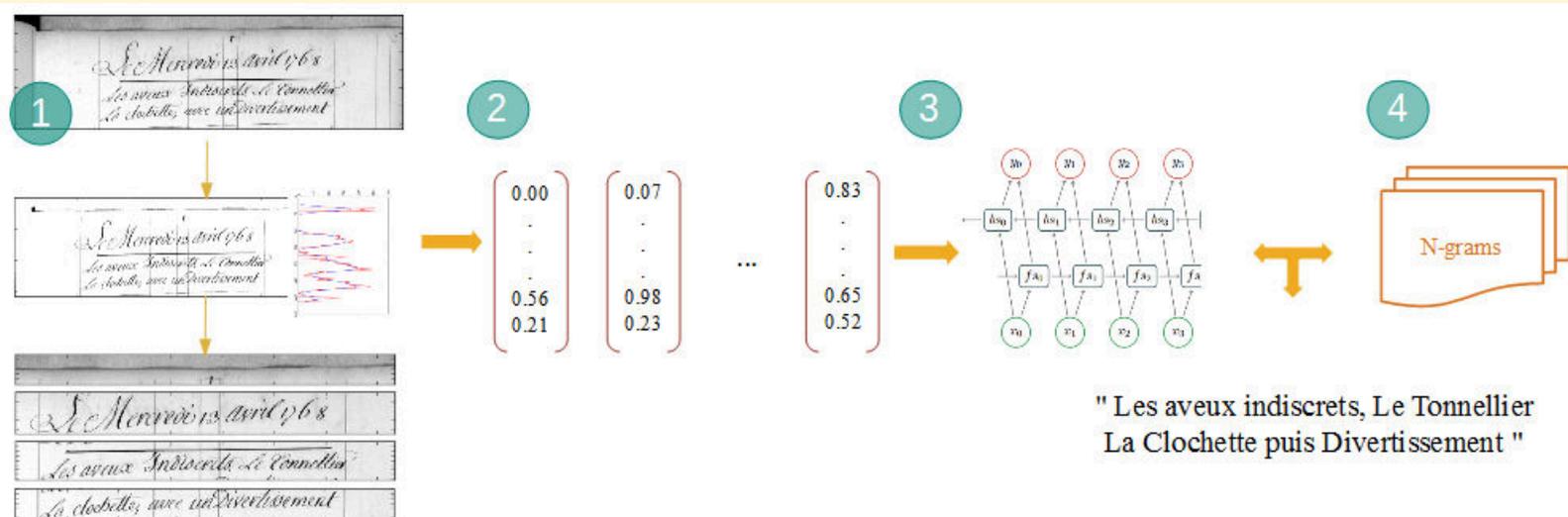
OBJECTIF ATTEINT !

Le défi du mois était de trouver 100 titres de pièces et les transcrire. Ouf, le défi est tout juste réalisé, on a eu chaud ! Bravo à vous, et n'hésitez pas à participer au prochain défi !

SCIENCES & TECHNO

Ce mois-ci, la rubrique est consacrée à la présentation des travaux d'Adeline Granet, doctorante au LS2N (Laboratoire des Sciences du Numérique de Nantes)

Adeline travaille à la conception de méthodes de reconnaissance d'écriture automatique. Si de nombreux travaux existent dans ce domaine, grâce à l'essor des capacités de calcul des ordinateurs et de méthodes issues de l'intelligence artificielle, les documents manuscrits anciens résistent encore et toujours à ces techniques !



L'image ci-dessus illustre l'approche adoptée par Adeline pour sa méthode de reconnaissance automatique des titres de pièces apparaissant dans les registres que vous retrouvez sur RECITAL.

À partir d'une image en couleur, les pixels de l'image sont classés pour identifier les zones d'encre (étape 1). Puis, ces pixels sont analysés en fonction de leur position, forme, etc., (étape 2) pour être utilisés par des systèmes de reconnaissance d'écriture tels que des réseaux de neurones (modèles de calcul pour l'apprentissage automatique fonctionnant sur le principe du cerveau humain) et des algorithmes statistiques comme les modèles de Markov (étape 3). Enfin, des techniques de traitement automatique du langage sont appliquées pour reconstruire le texte original (étape 4).

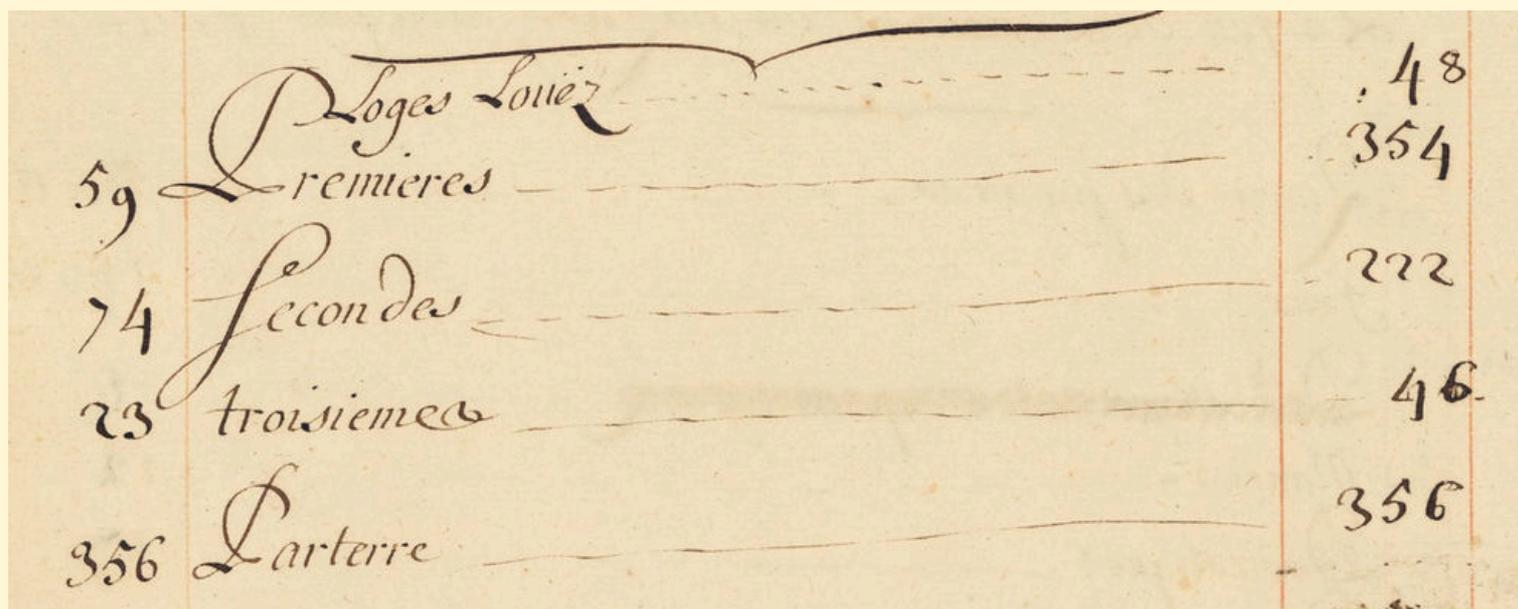
Ces travaux ont fait l'objet d'un article qui a été accepté à la conférence LREC 2018 !

LREC (Language, Resources and Evaluation Conferences) est une conférence internationale dans le domaine du traitement des langues. Elle s'attache à valoriser les travaux sur de nouvelles données, notamment sur la collecte de données via le crowdsourcing.

Notre article intitulé "Crowdsourcing-based Annotation of the Accounting Registers of the Italian Comedy" décrit comment le crowdsourcing peut améliorer les modèles utilisés par Adeline et vice versa !

PROCHAIN DÉFI

Aidez-nous à atteindre le prochain objectif !



A handwritten ledger on aged paper with a large green checkmark watermark in the background. The ledger has five rows with columns for descriptions and amounts. The entries are:

	Loges Louiez	48
59	Premieres	354
74	Secondes	222
23	troisiemes	46
356	L'arterre	356

Ce mois-ci, l'objectif est de vendre un maximum de billets ! Trouvez les ventes de billets pour transcrire les montants.

Combien allez-vous vendre de billets ?

Nous verrons cela dans le prochain numéro