




NEWS EE




A Digital Investigator for
Historical Newspapers



Dans le cadre de ce projet, des chercheurs en sciences humaines, des informaticiens et des bibliothécaires travaillent ensemble sur la presse historique conservée dans trois bibliothèques nationales (Autriche, Finlande, France).

 <https://www.univie.ac.at/newseye/>

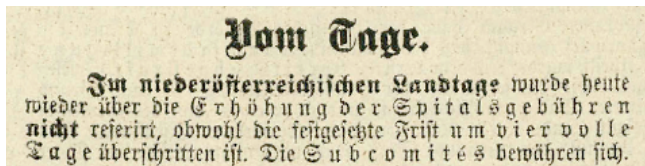
 www.newseye.eu  [@newseyeeu](https://twitter.com/newseyeeu)

Reconnaissance automatique du texte

Contact :

Roger Labahn

✉ roger.labahn@uni-rostock.de



Vom Tage.

Im niederösterreichischen Landtage wurde heute wieder über die Erhöhung der Spitalsgebühren nicht referirt, obwohl die festgesetzte Frist um vier volle Tage überschritten ist. Die Subcomités bewähren sich.

La reconnaissance automatique du texte fournit une transcription pour chaque ligne de texte.

La reconnaissance automatique du texte consiste à transcrire automatiquement les contenus textuels d'une image. C'est une étape fondamentale dans le flux de travail de NewsEye car elle fournit les données textuelles nécessaires pour être ensuite traitées dans des opérations de traitement automatique du langage naturel. Dans notre cas de figure, nous appliquons des modèles de reconnaissance à des images de

lignes de textes segmentées qui sont le produit d'un algorithme d'analyse de la mise en page. Ces modèles impliquent des réseaux de neurones artificiels profonds conçus pour des tâches d'étiquetage de séquences. De plus, les résultats de ces modèles incluent des probabilités appliquées à la reconnaissance de chaque caractère, qui peuvent être utilisées pour chercher dans les transcriptions au moyen de techniques de repérage de mots-clés (*keyword spotting*).