

# THÈSE DE DOCTORAT DE

NANTES UNIVERSITÉ

ÉCOLE DOCTORALE N° 641  
*Mathématiques et Sciences et Technologies  
de l'Information et de la Communication*  
Spécialité : *Informatique*

Par

**Léane JOURDAN**

## **Automatic Text Revision for Scientific Writing Assistance**

Révision automatique de texte pour l'aide à l'écriture scientifique

Thèse présentée et soutenue à Nantes, le 4 décembre 2025

Unité de recherche : UMR6004 – Laboratoire des Sciences et du Numérique de Nantes (LS2N)

### **Rapporteurs avant soutenance :**

Claire GARDENT Directrice de Recherche, CNRS/LORIA - Université de Lorraine  
Frederic BECHET Professeur des universités, Aix Marseille Université - LIS

### **Composition du Jury :**

Président :	Prénom NOM	Fonction et établissement d'exercice ( <i>à préciser après la soutenance</i> )
Examineurs :	Liana ERMAKOVA	Maîtres de conférences, Université de Bretagne Occidentale - HCTI
	Rachel BAWDEN	Chargé de recherche, Inria
	Marco KUHLMANN	Professor, Linköping University
Dir. de thèse :	Richard DUFOUR	Professeur des Universités, Nantes Université - LS2N
Co-dir. de thèse :	Florian BOUDIN	Chargé de recherche, HDR, Inria
Encadrant :	Nicolas HERNANDEZ	Maître de conférences, Nantes Université - LS2N

### **Invitée :**

Akiko AIZAWA Professor, National Institute of Informatics

**Titre :** Révision automatique de texte pour l'aide à l'écriture scientifique

**Mot clés :** révision de texte, assistance à l'écriture, articles scientifique, jeu de données

**Résumé :** Rédiger un article scientifique est une tâche difficile car il s'agit d'un genre littéraire avec ses propres conventions, ainsi de bonnes compétences d'écriture sont essentielles afin de transmettre les résultats de ses recherches. Au sein du processus d'écriture, la révision est une étape essentielle garantissant la clarté, la cohérence et le respect des normes académiques. Cette thèse s'intéresse à la révision de textes scientifiques en mettant l'accent sur le développement de ressources et de stratégies d'évaluation pour l'aide à l'écriture automatique. Nos contributions principales sont les suivantes : Afin d'étayer la recherche empirique, nous présentons CASIMIR, un grand jeu de données regroupant des révisions d'articles scientifiques, alignés au niveau des phrases et enrichis d'opérations d'édition et d'intentions

de révision inférées. Conscients des limites du travail au niveau de la phrase, nous proposons une définition de la tâche au niveau du paragraphe, soutenue par ParaRev, un jeu de données de paragraphes scientifiques révisés annotés avec des instructions de révision détaillées. Nos expériences révèlent que ces instructions adaptées au contexte permettent d'améliorer les révisions automatiques. Enfin, nous abordons les limites des mesures d'évaluation actuelles basées sur la similarité qui ne reflètent pas correctement la qualité des révisions de textes scientifiques. Nous explorons d'autres approches, notamment des mesures sans référence et l'évaluation LLM-as-a-judge, et constatons qu'une stratégie combinant des mesures basées sur les LLM et d'autres spécifiques à la tâche permet de saisir de manière plus fiable la qualité des révisions.

**Title:** Automatic Text Revision for Scientific Writing Assistance

**Keywords:** text revision, writing assistance, scientific articles, dataset

**Abstract:** Writing a scientific article is a challenging task as it is a highly codified genre, consequently proficiency in written communication is essential for effectively conveying research findings. In the writing process, text revision involves substantial modifications to improve clarity, coherence, and adherence to academic standards. This thesis investigates scientific text revision with a focus on the development of resources and evaluation strategies for automated writing assistance. Our main contributions are as follows: To support empirical research, we introduce CASIMIR, a large-scale dataset of scientific article revisions, aligned at the sentence level and enriched with edit operations and inferred revision intentions. Rec-

ognizing the limitations of sentence-level scope, we propose a paragraph-level definition of the task, supported by ParaRev, a dataset of revised scientific paragraphs annotated with fine-grained revision instructions. Our experiments reveal that those context-aware instructions improve automated revisions. Finally, we address the limitations of current similarity-based evaluation metrics which inadequately reflect the quality of scientific text revisions. We explore alternative approaches, including reference-free metrics and LLM-as-a-judge evaluation, finding that a hybrid strategy combining LLM-based and task-specific metrics most reliably captures revision quality.