



Avis de Soutenance

Monsieur Arthur HEMMER

Spécialité : Informatique et Applications

Soutiendra publiquement ses travaux de thèse intitulés

« Du Document à la Décision : Approches Neurosymboliques pour l'Extraction d'Informations à Partir de Documents Transactionnels »

dirigés par Monsieur Mickael COUSTATY

Soutenance prévue le vendredi 28 novembre 2025 à 15h00

Lieu : Université de La Rochelle

Salle : **Amphi Michel Crépeau**

3 avenue Albert Einstein

17031 La Rochelle

Composition du jury proposé

M. Mickael COUSTATY	La Rochelle Université	Directeur de thèse
M. Joseph CHAZALON	EPITA	Examineur
Mme Parisa KORDJAMSHIDI	Michigan State University	Examinatrice
M. Antoine DOUCET	La Rochelle Université	Examineur
M. Adam JATOWT	University of Innsbruck	Rapporteur
Mme Nicole VINCENT	Université Paris Cité	Rapporteure
M. Jean-Marc OGIER	CESI	Invité
M. Eric SIBONY	Shift Technology	Invité

Résumé :

Cette thèse porte sur l'amélioration de l'extraction d'informations à partir de documents en corrigeant les erreurs grâce à l'utilisation de connaissances préalables spécifiques au domaine. Elle se concentre sur les documents transactionnels tels que les factures, reçus et devis, pour lesquels nous formulons des contraintes symboliques permettant de vérifier l'exactitude des extractions. Ce mécanisme de validation sert alors de signal pour explorer des hypothèses d'extraction alternatives jusqu'à en identifier une valide. Les hypothèses alternatives proviennent à la fois de la reconnaissance optique de caractères et de l'étape d'extraction d'informations dans la chaîne de traitement documentaire. La validation est également appliquée aux modèles de bout en bout via la génération de données synthétiques, où les exemples invalides sont filtrés et les exemples restants, de haute qualité, sont utilisés pour affiner les modèles. Les résultats montrent que les approches neurosymboliques permettent d'obtenir une meilleure précision, en particulier dans des contextes frugaux, la qualité des contraintes jouant alors un rôle déterminant.