



Institut National de Statistique  
et d'Economie Appliquée



Centre des Etudes Doctorales  
Sciences, Ingénierie  
et Développement Durable

## Avis de soutenance de thèse de Doctorat

### Monsieur Idris OUMOUSSA

Doctorant au laboratoire de recherche

« Systèmes d'Informations, Systèmes Intelligents et Modélisation Mathématique »  
(SI2M)

*Spécialité : Systèmes d'Informations et Systèmes Intelligents*

Soutiendra publiquement sa thèse de Doctorat

Le jeudi 06 juin 2024 à 14h

à la salle de conférence de l'INSEA

Intitulé de la thèse

« **Business process-driven microservices identification :  
an NLP approach for supporting legacy system migration** »

Devant le jury composé de :

**Président :**

Pr. Adil KABBAJ, PES, INSEA-Rabat

**Directrice de thèse :**

Pr. Rajaa SAIDI, PES, INSEA-Rabat

**Membres du jury :**

Pr. Ahmed ZELLOU, PES, ENSIAS-Rabat

Pr. Driss ALLAKI, PH, INPT-Rabat

Pr. Maryam RADGUI, PH, INSEA – Rabat

Pr. Naouel MOHA Professeur à l'École de Technologie Supérieure (ÉTS)  
Université du Québec



**INSEA**  
Institut National de  
Statistique et d'Économie  
Appliquée  
**CEDOC-SIDD**

## *Les Résumés de la thèse (F1)*

Date : 01/04/2024

Réservé à l'administration

N° de thèse : .....

Nom : Oumoussa

Prénom : Idris

---

## *Résumé*

---

Les microservices se sont avérés être une alternative précieuse à de nombreuses technologies existantes, permettant la décomposition d'applications monolithiques en composants/modules de granularité fine, hautement cohésifs et faiblement couplés. Toutefois, l'identification des microservices demeure un défi majeur pour le succès d'une activité de migration. Cette identification implique la récupération, parmi les systèmes existants, des fonctionnalités de service potentielles ainsi que des artefacts réutilisables qui peuvent être regroupés tout en conservant une logique métier significative. Le défi consiste à identifier les services potentiels, issus des systèmes existants, qui peuvent être développés de manière rentable, adaptés à la réutilisation, faciles à entretenir, tout en permettant une personnalisation adéquate des applications finales grâce à une sélection et une orchestration appropriées des services.

Cette thèse soutient la migration des systèmes légataires vers les microservices en (1) analysant les travaux existants concernant l'identification de microservices dans le contexte académique et industriel tout en examinant les objectifs de recherche, les obstacles existants, et les nouveaux défis qui se profilent, et (2) en proposant une approche d'identification automatique des microservices à partir d'un ensemble de processus métiers. Notre approche descendante, basée sur le traitement du langage naturel pour déterminer la similarité entre les activités impliquées dans ces processus métiers, s'appuie sur un regroupement rigoureux pour rassembler les activités en microservices. Afin de mettre en lumière notre approche et de prouver son efficacité, nous l'avons expérimentée sur deux études de cas. Les résultats expérimentaux mettent en évidence comment les dépendances entre les activités, extraites des spécifications des processus métiers, permettent de générer des microservices de qualité, proposant ainsi des améliorations considérables par rapport à d'autres approches citées dans la littérature.

**Mots clés:** Microservices, Processus métier, BPMN, Décomposition des applications monolithiques, NLP, Clustering.

## *Abstract*

---

Microservices have proved to be a valuable alternative to many existing technologies, enabling monolithic applications to be broken down into highly cohesive, loosely coupled components/modules of fine granularity. However, identifying microservices remains a major challenge, crucial to the success of these migrations. This involves identifying potential service functionalities within existing systems, as well as reusable artefacts that can be grouped together while retaining valuable business logic. The challenge is to identify potential services from legacy systems that can be developed cost-effectively, suitable for reuse, easy to maintain, while allowing adequate customisation of the final applications through appropriate service selection and orchestration.

This thesis supports the migration of legacy systems to microservices by (1) analysing the state of microservices identification practices in academia and industry while examining research objectives, existing barriers, and emerging challenges, and (2) proposing an approach for the automatic identification of microservices based on a set of Business Processes (BP). Our top-down approach based on Natural Language Processing (NLP) to determine the similarity between activities involved in these business processes, relies on rigorous clustering to group activities into microservices. In order to highlight the approach and prove its effectiveness, we opted for two concrete case studies. The experimental results show how dependencies between activities, extracted from business process specifications, can be used to generate high-quality microservices compared to other approaches cited in the literature.

**Keywords:** Microservices, Business processes, BPMN, Decomposition of monolithic applications, NLP, Clustering.



## ملخص

أثبتت الخدمات الصغيرة أنها بديل قيّم للعديد من التقنيات القائمة، حيث تُمكن من تفكيك التطبيقات الضخمة إلى مكونات/وحدات فرعية صغيرة الحجم ذات تماسك عالي وارتباط منخفض. ومع ذلك، يبقى تحديد الخدمات الصغيرة تحدياً رئيسياً لنجاح عمليات الترحيل هذه. يتضمن ذلك تحديد وظائف الخدمة المحتملة داخل الأنظمة القائمة، وكذلك تعيين العناصر القابلة لإعادة الاستخدام التي يمكن تجميعها مع الاحتفاظ بمنطق الأعمال القيّم. يكمن التحدي في تحديد الخدمات المحتملة من الأنظمة القديمة التي يمكن تطويرها بتكلفة فعالة، ومناسبة لإعادة الاستخدام، وسهلة الصيانة، مع السماح بتخصيص كافٍ للتطبيقات النهائية من خلال اختيار الخدمات المناسبة وتنظيمها.

تدعم هذه الأطروحة عملية ترحيل الأنظمة القديمة نحو الخدمات الصغيرة عن طريق (1) تحليل تقنيات تحديد الخدمات الصغيرة في المجال الأكاديمي والصناعي مع دراسة أهداف البحث في هذا المجال عامة، العوائق القائمة به و التحديات الناشئة فيه، وكذلك (2) اقتراح نهج تلقائي لتحديد الخدمات الصغيرة بناءً على مجموعة من العمليات التجارية. يعتمد نهجنا على معالجة اللغة الطبيعية لتحديد التشابه بين أنشطة العمليات التجارية، و يستند على تجميع صارم للأنشطة في خدمات صغيرة. من أجل إبراز نهجنا وإثبات فعاليته، اخترنا دراسة حالتين عمليتين. تُظهر النتائج التجريبية كيف يمكن استخدام التبعيات بين الأنشطة، التي تم استخراجها من مواصفات عمليات الأعمال التجارية، لإنتاج خدمات صغيرة عالية الجودة مقارنة بالنهج الأخرى المذكورة في الأدبيات.

**الكلمات الرئيسية:** الخدمات المصغرة، العمليات التجارية، BPMN، تحليل التطبيقات المتجانسة، معالجة اللغة الطبيعية، التجميع.