

Compte rendu : GenAI

Find an internship abroad

Jessy Latmi

Résumé

Ce rapport décrit la mise en place d'un workflow d'automatisation pour la recherche de stages à l'étranger, utilisant n8n et des modèles de langage. Le workflow permet d'entrer des villes et technologies, de rechercher des offres sur le web, d'extraire automatiquement des informations clés telles que dates, compétences et salaires, et de transmettre les résultats via un webhook Discord.

Sommaire

1	Introduction	3
2	Workflow d'automatisation	4
2.1	Image du workflow	4
2.2	Explication du workflow	4
2.3	Exemple de résultat final	4
2.4	Exemple de résultat brut	5
3	Analyse et perspectives	6

1 Introduction

La recherche de stages à l'étranger constitue un enjeu majeur dans le cadre de la formation d'ingénieur. Cependant, cette tâche peut s'avérer fastidieuse : il est nécessaire de consulter de multiples sites d'offres, de filtrer des informations pertinentes et de suivre les mises à jour en continu.

Dans ce contexte, l'automatisation apparaît comme une solution pour gagner du temps et augmenter l'efficacité de la recherche. L'émergence des plateformes d'automatisation, comme n8n, et des modèles de langage avancés (LLM) permet de concevoir des workflows capables de collecter, filtrer et traiter des informations complexes provenant de sources variées, sans nécessiter un développement complet.

Ce rapport décrit la conception et la mise en œuvre d'un workflow d'automatisation pour la recherche de stages à l'étranger. Le système présenté combine n8n pour l'orchestration des tâches et un LLM pour l'extraction automatique des informations clés, telles que les dates de stage, les compétences requises et les fourchettes de salaire. L'objectif est de démontrer comment ces outils peuvent être utilisés conjointement pour simplifier et structurer un processus habituellement manuel et répétitif.

2 Workflow d'automatisation

2.1 Image du workflow

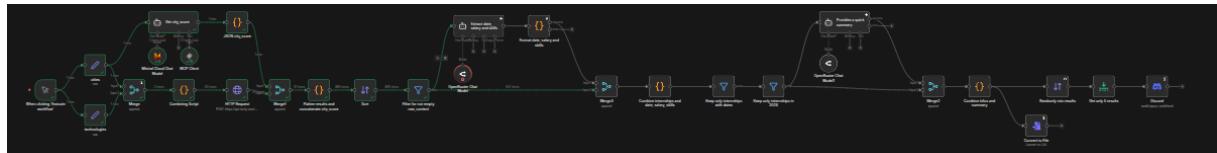


FIGURE 1 – N8N Workflow

2.2 Explication du workflow

1. Entrée manuelle d'une liste de villes et de technologies
 2. Combinaison des villes et des technologies
 3. Recherche de 15 résultats web pour chaque combinaison (via l'API Tavily)
 4. Ajout d'un score correspondant à la ville pour chaque requête
 5. Extraction des dates, du salaire et des compétences par un LLM
 6. Filtrer pour ne garder que les offres avec des dates en 2026
 7. Récupérer 3 résultats au hasard
 8. Envoyer les résultats par Discord

2.3 Exemple de résultat final

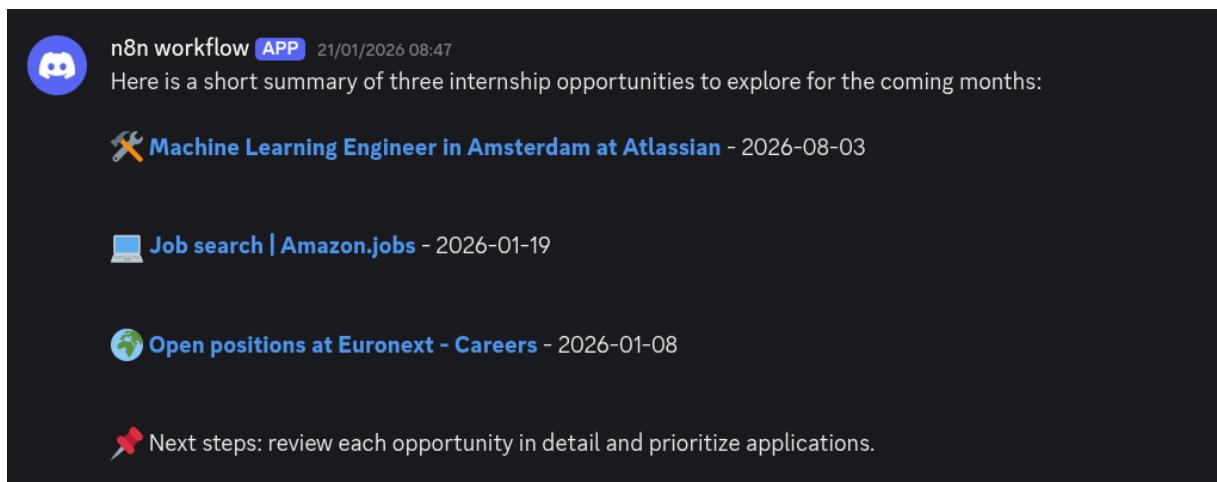


FIGURE 2 – Webhook Discord

2.4 Exemple de résultat brut

Vous retrouverez un JSON plus complet dans le repository Gitlab.

```
{  
    "index": 6,  
    "output": "The company is AstraZeneca, focused on transforming healthcare through effective data-driven teamwork. Internships in Barcelona need basic Python skills for analyzing data.",  
    "answer": "Internships in Barcelona for Python careers include roles like Data Analyst and Junior Python Developer. These positions require basic Python skills and often involve data analysis. Opportunities are available in both tech and gaming sectors.",  
    "content": "Join a talented global team that's powering AstraZeneca to better serve patients every day. When you join us, you'll be part of a team that embraces digital technology and data to transform the way we work and the work we do. There are many things I enjoy when working at AstraZeneca, mainly the Speak up culture, the great colleagues that are in my teams, the great products that AstraZeneca provides to our patients and the challenging conversations I have around our medicines. There are many things I enjoy when working at AstraZeneca, mainly the Speak up culture, the great colleagues that are in my teams, the great products that AstraZeneca provides to our patients and the challenging conversations I have around our medicines. Working at AstraZeneca has impacted my life in such a positive way. At AstraZeneca, our purpose is to help patients all over the world by delivering life-changing medicines as one collaborative team.",  
    "city_score": 96.94,  
    "url": "https://careers.astrazeneca.com/job/barcelona/data-and-ai-engineer-intern/7684/90680913728",  
    "title": "Data & AI Engineer Intern",  
    "date": "2026-01-17",  
    "salary": null,  
    "skills": [  
        "Python",  
        "data analysis",  
        "data wrangling",  
        "basic programming"  
    ]  
}
```

3 Analyse et perspectives

L'implémentation de ce workflow a permis de mettre en évidence plusieurs points clés sur l'usage combiné de n8n et des modèles de langage pour l'automatisation de tâches concrètes :

- **Efficacité et gain de temps :** L'automatisation permet de générer rapidement un ensemble d'offres pertinentes, d'extraire les informations essentielles et de les transmettre sur Discord. Cela réduit considérablement le temps nécessaire pour effectuer ces recherches manuellement.
- **Fiabilité des informations :** L'utilisation de LLM pour l'extraction des données présente une variabilité inévitable. Les résultats dépendent de la qualité des prompts et de la structuration du contenu source. Certaines informations, comme les salaires ou les compétences spécifiques, peuvent être manquantes ou mal interprétées.
- **Complexité technique du workflow :** Malgré l'aspect no-code de n8n, la création d'un workflow avancé a nécessité l'écriture de scripts JavaScript pour gérer certaines transformations et filtres de données. L'absence de multi-input sur certains blocs a également contraint à concevoir des solutions détournées pour traiter les combinaisons villes/technologies.
- **Perspectives d'amélioration :** Plusieurs axes peuvent être envisagés pour rendre le workflow plus robuste et polyvalent :
 1. Ajouter un filtrage plus avancé, basé sur les compétences, les salaires ou les durées de stage.
 2. Utiliser des modèles de langage spécialisés ou des techniques de post-traitement pour améliorer la précision des informations extraites.
 3. Spécialiser le workflow à des plateformes de publication d'offres afin d'obtenir des informations plus cohérentes entre-elles et automatiser la mise à jour quotidienne des résultats.

En résumé, ce projet illustre le potentiel des outils no-code et des modèles de langage pour automatiser des processus complexes. Il met également en lumière les limites actuelles et ouvre la voie à des améliorations qui pourraient rendre ce type de workflow plus fiable et plus complet pour la recherche de stages à l'étranger. Pour l'heure, si n8n est un outil intéressant d'automatisation, la pertinence d'utiliser un LLM plutôt que du web scraping classique afin d'extraire des informations est à remettre en question.