

Stage voor de opleiding Master Informatica

Titel: ML of Software Engineering stage

Gegevens bedrijf:

Naam: ML6

Tel: 09 265 95 50

Contactpersoon: Amse D'Oosterlinck
mailadres: jobs@ml6.eu

Adres waar de student zal werken: Esplanade Oscar Van De Voorde
1, 9000 Gent

Kort (min 90 uur) of lang (180 uur): **kort / lang**

Korte beschrijving van de opdracht: keuze uit 9 domeinen

Open Internship domains:

Information Management

The Information Management unit powers the ML6 teams with effective information. In order to this, we develop a series of internal tools (e.g. dashboards, applications, etc). We offer you an opportunity to work hands-on on applications, be it backend and/or frontend. We are looking for candidates that have worked with mainstream backend (preference: Python) and/or frontend languages (preference: React).

GenAI NLP

LLMs like GPT promise to fundamentally change knowledge work. Many challenges remain, however, related to alignment, security, truthfulness, bias, performance, latency, cost and others. Work on one of these problems and help move forward the field towards truly useful, robust and measurable large language model applications! Potential topics include local language LLMs, domain-specific fine-tuning, performance benchmarking of LLMs, tabular data understanding, building an internal InTRAGnet bot...

GenAI Vision

Image generation models like Stable Diffusion are revolutionizing the creative industries and beyond by sharply increasing the reach and productivity of designers and allowing lay people to create prototypes in

minutes. Techniques like Dreambooth and ControlNet further allow for the guidance of generative models in clever ways. Help us push the boundaries of AI image generation by researching new ways of guiding image generation and develop domain-specific use cases. Potential topics include domain-specific fine-tuning (e.g. food photography, fashion...), ControlNet guidance, Gligen...

Digital Twins - Industry 4.0 🏭

This domain focuses on using AI in the context of Industry 4.0. Think about digital twins, predictive maintenance, root cause analysis, process optimisation, process steering. We also see value for use cases where labeling data is expensive or difficult to do, which we tackle by using segment anything model, active learning or surrogate models. Potential subtopics include PINNs, Digital Office, Azure IoT...

Fondant - open source 📦

Foundation models learn new knowledge and skills through cleverly prepared data. Help us push the boundaries of data creation, curation and augmentation by researching new functionalities and developing components and pipelines for the Fondant open source framework developed by ML6. Potential subtopics include bias removal, knowledge domain-based filtering, synthetic data generation, topic-based fine-tuning, distillation, Python (/TypeScript) software development...

AI in Bio 🧬

Biology is intrinsically complex and diverse. Despite, decades of research, nature still holds many secrets. Today, there is an opportunity for AI to support the experts: mapping experiment input and output over very complex biological functions. Potential topics include Drug Discovery and Omics analysis.

Energy ⚡

AI can optimize resource management, enhance grid efficiency, and accelerate the transition to clean energy. This intersection holds the key to a sustainable and greener future, making it a crucial avenue for innovation and progress.

Potential internship topics within the energy domain include leveraging and combining open-data forecasts to improve anticipated renewable energy generation, balancing the net with AI, closing the gap between energy finance and actual supply-demand by deepdiving into forecasted energy

trading and optimising residential energy consumption. For the last one you could even get to use your own smart meter data!

Ops 🖱️

Operations is a term subsuming DevOps and MLOps. DevOps refers to a set of standards, practices and tooling that enable fast/high quality delivery and continuous improvements. MLOps extends DevOps to ML systems, which come with a set of challenges related to the training, deployment, monitoring and tracking of models.

Hybrid Solutions 💡

For some of our most frequently solved challenges we offer hybrid solutions. These projects are based on a shared codebase and internships are centered on extending the shared codebase based on latest technical developments. Applications range from process steering in manufacturing to computer vision diagnostics and entity recognition in legal documents. Potential topics are domain adaptation, zero shot learning and more.

Technologieën die aan bod zullen komen:

- Familiarity with statistical analysis languages and tools like Python, SQL
- Excellent verbal and written communication in English
- You are currently pursuing a degree in computer science or related field
- A first experience with Tensorflow/ Pytorch and cloud deployment is considered a plus
- Cloud (GCP, AWS, Azure)