

# Online Learning NMT

BNP Paribas CIB AI Lab develops its own automated translation engine: come & help us build & improve one of the most iconic AI tools today! The tool is available and used by more than 200 000 employees within the group providing high quality translation that can be fine-tuned to their needs.

Since 2015, almost all the translation competition winners used neural machine translation algorithms, which became the state-of-the-art in almost all the tasks of the field.

The data science team developed a deep expertise on these state-of-the-art architectures [1][2]. The tool is widely adopted and more and more employees require models meeting their custom needs.

During the internship, the student will explore the **development of a new feature based on Online Learning** [4]. Using online learning and human correction, the idea is to **automatically update the translation model without degrading the initial performances**. The intern will challenge the most advanced pipeline that are using **Dynamic Knowledge Distillation** [5].

This could also be the opportunity to practice some backend & frontend development, delivering a first POC around Online Learning from users input and improving the Translate user experience.

This work will explore the boundaries of **academic research** to build potential new applications within the BNP Paribas ecosystem.

## References:

- [1] Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation, Wu & AL, CoRR, vol. abs/1609.08144 (2016)
- [2] Attention Is All You Need, Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin, NIPS 2017
- [3] Six Challenges for Neural Machine Translation, Koehn P. & Knowles R., 2017 In Proceedings of the First Workshop on Neural Machine Translation.
- [4] Findings of the first shared task on lifelong learning machine translation Barrault, L., Biesialska, M. M., Ruiz Costa-Jussà, M., Bougares, F., & Galibert, O. (2020). Findings of the first shared task on lifelong learning machine translation. In EMNLP 2020,
- [5] Continual Learning for Neural Machine Translation, Yue Cao, Hao-Ran Wei, Boxing Chen, XiaoJun Wan

## Practical Information:

**Supervisor:** MATHEY Alexis, Data Scientist, [alexis.mathey@bnpparibas.com](mailto:alexis.mathey@bnpparibas.com)

**Location:** Grands Moulins de Pantin, 9 Rue du Débarcadère, 93500 Pantin (Possible partial work from home)

**Compensation:** Competitive salary

**How to apply:** Please submit your application at [paris.cib.analytics.consulting.careers@bnpparibas.com](mailto:paris.cib.analytics.consulting.careers@bnpparibas.com)

