



# Free Format Resume Entity Recognition

- Summer Internship at Radix -

## About Radix

Radix is an Artificial Intelligence solutions provider based in Brussels. Together with our clients, we deliver AI solutions that have a positive impact on the world.

Our mission is to help our clients grow and to improve people's lives. To deliver on our promise, we focus on the combination of three elements:

- Business: creating high value for our clients, their employees and society as a whole
- People: delivering technology that can truly assist people, not replace them
- AI: automating intelligence to achieve the highest impact

Founded in January 2018, Radix is now a team of 25 people that continues to grow. We work with ambitious clients like GSK, Brussels Airport, Atlas Copco, Flanders Investment and Trade, VDAB, Belga, Macadam, House of HR and more.

Visit our website <https://radix.ai> for more information.



## The internship

As of this year, Radix will host several summer internships. Such an internship shows you, among other things, how your learning material is applied in the field and what the best practices are when working on a project, this while being supervised by field experts. Aside from that, it allows you to get a glance of how a growing startup operates from the inside and how other fields such as *Marketing* and *Sales* contribute to this.

### About the project

Most ML models require plenty of data to train on. Although the magnitude of collected data continues to increase steadily, this data is often very noisy or lacks proper labels in order to be used during model training. One way to solve this problem is to create another model with the only task to clean and structure the data. During this internship, you'll create such a model that extracts required entities from free-format resumes. More specifically, you'll detect entities such as contact information, fields of study, ... from PDFs and classify these accordingly.

### What you will learn

In order to succeed at this internship, you'll learn and apply various aspects of traditional machine learning. First, you'll learn to acquire, modify and augment resume data used to train our model. On this data, a computer vision model is applied to detect the parts of interest for every resume. Given these extracted segments, you'll train a *Named Entity Recognition* algorithm to extract and classify entities such as contact information, fields of study, or job titles.

## Your profile

- You are in your third Bachelor or first Master in Computer Science or a related field
- You are fluent in English, both written and verbally
- You have strong analytical skills and are familiar with the basics of machine learning
- You know your way around in Python
- Familiarity with statistical Python tools such as Spacy, TensorFlow, or Torch is a plus
- Previous experience on a related ML project is a big plus
- You are eager to learn and are open to be challenged