We research, investigate and develop new technologies for digital audiovisual media. Together with our shareholders ARD, ZDF, Deutschlandradio, ORF and SRG/SSR and international partners, we work on innovations and practical solutions in the rapidly changing media environment.

Do you love to experiment? Do you like to work on new technologies and want to bring in your own ideas? Then join our team of experts and write your

**Bachelor Thesis**

*on the topic „Data aggregation for Natural Language Processing for the German language“*

Automatic processing and understanding of human written texts have become essential in many branches of media technology. Natural Language Processing (NLP) technologies alongside deep learning offer the ability to decipher, translate and process texts. These technologies have been developed rapidly, however mainly in the English language. One of the major challenges in the german media industry today is to transpose and implement these advances to the German language. You will design and implement a pipeline for data gathering from news articles, including data preparation for Natural Language Processing tasks.

What you bring:

- Field of study: computer science, informatics, engineering or comparable field of study
- Basic knowledge, ideally first assignments, in Python
- Ideally basic knowledge around databases, SQL, Natural Language Processing
- Interest in machine learning
- Motivation and ability to work independently as well as in a team
- Good communication skills in German and English

Your perspective at IRT:

You can work on exciting and challenging projects concerning current and future topics of media in an internationally renowned research and competence centre. We work in a highly motivated team with horizontal organisation structures. Flexible working hours as well as an attractive set of social benefits complete our offer to you. You can find further information at www.irt.de.