



Internship offer (M/F)
Reference: DS_PPfact_Intern_Jan21

Research topics	Fact-Checking, NLP, Data integration
Department	Data Science
Publication date	25-1-2021
Start date	ASAP
Duration and conditions	Six months internship (Stipend + Lunch coupons)
Place of the internship	BIOT Sophia Antipolis (teleworking is possible during pandemic period upon authorisation from supervisor and from FRANCE only).

Description

Our group has been working on methods for automatically fact checking text based on structured resulting in several top-tier conference papers and multiple industrial collaboration including one with Google.org for COVID-19 claims. Our most recent system (<https://coronacheck.eurecom.fr>) focuses on how to computationally verify claims about the Coronavirus. This is motivated by a spread of misinformation with regards to the Coronavirus with implications ranging from "funny but harmless" to "extremely dangerous" (e.g., when trying out certain proposals for self-medication). The current version verifies statistical claims about the spread of Coronavirus, by translating text claims into SQL queries on data from sources such as WHO or CDC.

The system, in seven languages, already had north of 15,000 users (from 100+ countries) as well as press coverage by several newspapers (see <https://coronacheck.eurecom.fr/en/press> for details).

We have plans for multiple research directions which could be realized in an internship. This research would immediately result in extensions and improvements for our running system. They are directly motivated by feedback and queries from our growing user base.

A superset of research directions we aim to address is the following:

1) Our current Web interface returns can be extended by generating a short natural language report that summarizes arguments and assumptions under which certain verification results hold. This requires research on "explainable fact checking".

2) Our current system issues potentially many candidate SQL queries that may translate a given input claim. This works well as long as those queries execute on small data sets. While that applies to the data we're currently using, there have been various new data sets (e.g., data with details on vaccine) which are larger and heterogeneous in many aspects. Furthermore, we want to go from the verification of single claims to the verification of entire Web documents. Scaling up to such scenarios requires research on optimizing batch processing of candidate queries as well as on prioritizing processing of sets of candidate queries, based on their probability to be relevant for a given verification.

EURECOM, dans le cadre de son plan d'égalité hommes/femmes et de ses actions positives en faveur de la mixité des emplois, encourage particulièrement les candidatures féminines pour les postes d'ingénieurs et de chercheurs.



3) Finally, we observe claims in our logs that go beyond pure statistical claims (e.g., claims about whether specific types of medicine are helpful against the Coronavirus). We want to extend the coverage to those claims as well. We are currently already working towards doing that, leveraging our recent work, done in collaboration with Google, on matching entries from the claim reviews corpus.

Requirements

- Education Level / Degree : this is a final internship for master students
- Field / specialty: n.a.
- Technologies: ML, web development, unix/linux shell
- Languages / systems: python
- Other skills / specialties: knowledge of NLP (language models, transformers) is a plus
- Other important elements: drive to publish research papers is a plus

Application

The application must include:

- Detailed curriculum,
- Motivation letter of at most one page,
- List of exams and grades during master

Applications should be submitted by e-mail to secretariat@eurecom.fr with the reference: DS_PP_fact_Intern_Jan21

Important Dates: Screening will start immediately.

About EURECOM

EURECOM is a graduate school and a research center in communication systems located in Sophia Antipolis, a vibrant science park on the French Riviera. EURECOM is ranked among the world's top universities in the QS World University Rankings® 2019, considered one of the world's strongest universities in Computer Science & Information Systems and ranked 551/600 worldwide.

Organized as an Economic Interest Group (kind of consortium), EURECOM brings together in its consortium prestigious universities such as the schools from the Institut Mines Télécom group (Télécom Paris, Imt Atlantique, Télécom SudParis, etc.), Aalto University (Helsinki), Politecnico di Torino, Technische Universität München (TUM), Norwegian University of Science and Technology (NTNU), Chalmers University of Technology (Sweden), and Czech Technical University in Prague (CTU), ITMO University (St Petersburg), as well as industry members such as BMW Group, IABG, Orange, SAP, NortonLifeLock and the Principality of Monaco as an institutional member.

EURECOM has developed its expertise around three major fields: Digital Security, Data Science and Communication Systems. EURECOM is particularly active in research in its areas of excellence while also training a large number of doctoral candidates. Its contractual research, in which its industrial members actively participate, is widely recognized in Europe and contributes largely to its budget. Its strong links with various industries has enabled EURECOM, with the Institut Mines Télécom, to obtain the Carnot label, a label granted to research organizations which put partnership research at the heart of their strategy.

EURECOM, dans le cadre de son plan d'égalité hommes/femmes et de ses actions positives en faveur de la mixité des emplois, encourage particulièrement les candidatures féminines pour les postes d'ingénieurs et de chercheurs.