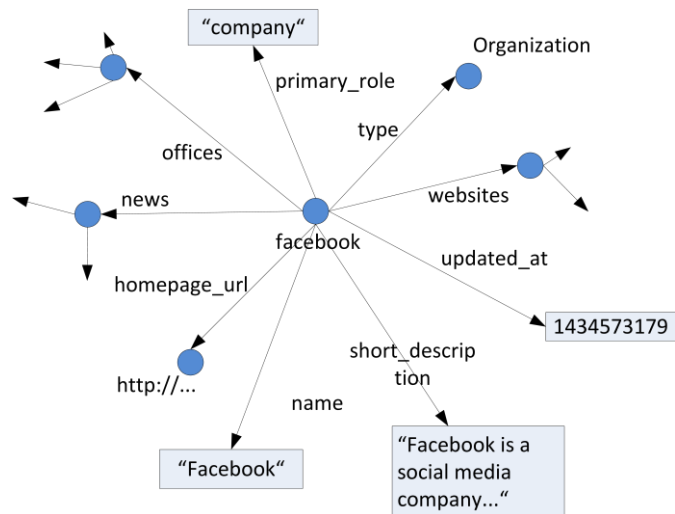


09.01.2020

Call for Bachelor/Master Thesis (in English or German) “How Do Successful Startups Look Like? Predicting the Success of Startups and Tech Companies”

What is the topic?

CrunchBase is an online platform providing information about startups and technology companies, including related entities such as the products they sell, key people they employ, and investments they made and received. In our research group, an RDF data set containing the Crunchbase data has been generated (see [1,2]). This data set has been used for data integration with financial data sources to evaluate the performance of particular companies and for monitoring news to find statements that are not in Crunchbase yet.



In 2018, a newer version of the data set with much higher coverage of the available data had been generated. The focus of this thesis is to use this data set for machine-learning prediction tasks. Approaches should be developed and evaluated which can predict the success of startups (e.g., which investors became interested in which companies, what was the investment amount, ...). The precise machine learning setup and the definition of “success” of companies will be fixed at the beginning of the thesis period.

Note that the data is available in RDF (nt files).

This thesis requires that the applicant has skills in machine learning and data mining.

[1] <https://zenodo.org/record/3270905>

[2] http://dbis.informatik.uni-freiburg.de/content/team/faerber/papers/CrunchBaseWrapper_SWJ2017.pdf

Which prerequisites should you have?

- Basic knowledge about RDF, ability to write simple SPARQL queries.
- Ability to work on data analysis and machine learning.

Keywords: Semantic Web, RDF, machine learning, supervised machine learning, startups.

Contact:
Dr.-Ing. Michael Färber
michael.farber@kit.edu